OIPE JC68
JUL 0 8 2003
PATENT & TRADEMARK OFFICE

UNITED STATES PATENT AND TRADEMARK OFFICE

I, Alan John SPARROW

translator to RWS Group plc, of Europa House, Marsham Way, Gerrards Cross, Buckinghamshire, England declare;

1. That I am a citizen of the United Kingdom of Great Britain and Northern Ireland.

2. That I am well acquainted with the German and English languages.

3. That the attached is, to the best of my knowledge and belief, a true translation into the English language of the accompanying copy of the specification filed with the application for a patent in Germany on 1 March 1999 under the number 199 08 837.3 and the official certificate attached hereto.

4. That I believe that all statements made herein of my own knowledge are true and that all statements made on information and belief are true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the patent application in the United States of America or any patent issuing thereon.

AJ Sparrow

For and on behalf of RWS Group plc

The 27th day of March 2003

FEDERAL REPUBLIC OF GERMANY

Certificate

BASF Aktiengesellschaft

of

Ludwigshafen/Germany

have filed a Patent Application under the title:

"Novel poly(ADP-ribose) polymerase genes"

on 1 March 1999 at the German Patent and Trademark Office.

The attached documents are a correct and accurate reproduction of the original submission for this Patent Application.

The German Patent and Trademark Office has for the time being given the Application the symbols C 12 N, C 12 Q and A 01 K of the International Patent Classification.

Munich, 1 June 1999

German Patent and Trademark Office

The President

pp

Seiler

File No: 199 08 837.3

Novel poly(ADP-ribose) polymerase genes

The present invention relates to novel poly(ADP-ribose) polymerase (PARP) genes and to the proteins derived therefrom; antibodies with specificity for the novel proteins; pharmaceutical and gene therapy compositions which comprise products according to the invention; methods for the analytical determination of the proteins and nucleic acids according to the invention; methods for identifying effectors or binding partners of the proteins according to the invention; methods for determining the activity of such effectors and use thereof for the diagnosis or therapy of pathological states.

In 1966, Chambon and co-workers discovered a 116 kDA [sic] enzyme that was characterized in detail in subsequent years and is now called PARP (EC 2.4.2.30) (poly(adenosine-5'-diphosphoribose) polymerase), PARS (poly(adenosine-5'-diphosphoribose) synthase) or ADPRT (adenosine-5'-diphosphoribose transferase). This enzyme has to date been unique in its activity, which is described below. It is referred to as PARP1 below to avoid ambiguity.

The primary physiological function of PARP 1 appears to be its involvement in a complex repair mechanism which cells have developed to repair DNA strand breaks. The primary cellular response to a DNA strand break appears moreover to consist of PARP1-catalyzed synthesis of poly(ADP-ribose) from NAD^+ (cf. De Murcia, G. et al. (1994) TIBS, 19, 172).

PARP 1 has a modular molecular structure. Three main functional elements have been identified to date: an N-terminal 46kDa DNA binding domain; a central 22kDa automodification domain to which poly(ADP-ribose) becomes attached, with the PARP 1 enzyme activity decreasing with increasing elongation; and a C-terminal 54 kDa NAD^+ binding domain. A leucine zipper region has been found within the automodification domain, indicating possible protein-protein interactions, only in the PARP from Drosophila. All PARPs known to date are presumably active as homodimers.

The high degree of organization of the molecule is reflected again in the strong conservation of the amino acid sequence. Thus, 62% conservation of the amino acid sequence has been found for PARP 1 from humans, mice, cattle and chickens. There are greater structural differences from the PARP from Drosophila. The individual domains themselves in turn have clusters of increased conservation. Thus, the DNA binding region contains two so-called zinc fingers as subdomains (comprising motifs of the type

$CX_2CX_{28/30}HX_2C$), which are involved in the Zn^{2+}-dependent recognition of DNA single strand breaks or single-stranded DNA overhangs (e.g. at the chromosome ends, the telomeres). The C-terminal catalytic domain comprises a block of about 50 amino acids (residues 859-908), which is 100% conserved among vertebrates (PARP "signature"). This block binds the natural substrate NAD^+ and thus governs the synthesis of poly(ADP-ribose) (cf. de Murcia, loc.cit.). The GX_3GKG motif in particular is characteristic of PARPs in this block.

The beneficial function described above contrasts with a pathological one in numerous diseases (stroke, myocardial infarct, sepsis etc.). PARP is involved in cell death resulting from ischemia of the brain (Choi, D.W., (1997) Nature Medicine, 3, 10, 1073), of the myocardium (Zingarelli, B., et al (1997), Cardiovascular Research, 36, 205) and of the eye (Lam, T.T. (1997), Res. Comm. in Molecular Pathology and Pharmacology, 95, 3, 241). PARP activation induced by inflammatory mediators has also been observed in septic shock (Szabo, C., et al. (1997), Journal of Clinical Investigation, 100, 3, 723). In these cases, activation of PARP is accompanied by extensive consumption of NAD^+. Since four moles of ATP are consumed for the biosynthesis of one mole of NAD^+, the cellular energy supply decreases drasticallly. The consequence is cell death.

PARP1 inhibitors described in the abovementioned specialist literature are nicotinamide and 3-aminobenzamide. 3,4-Dihydro-5[4-1(1-piperidinyl)butoxy]-1(2H)-isoquinolone [sic] is disclosed by Takahashi, K., et al (1997), Journal of Cerebral Blood Flow and Metabolism 17, 1137. Further inhibitors are described, for example, in Banasik, M., et al. (1992) J. Biol. Chem., 267, 3, 1569 and Griffin, R.J., et al. (1995), Anti-Cancer Drug Design, 10, 507.

High molecular weight binding partners described for human PARP1 include the base excision repair (BER) protein XRCC1 (X-ray repair cross-complementing 1) which binds via a zinc finger motif and a BRCT (BRCA1 C terminus) module (amino acids 372-524) (Masson, M., et al., (1998) Molecular and Cellular Biology, 18,6, 3563).

It is an object of the present invention, because of the diverse physiological and pathological functions of PARP, to provide novel PARP homologs. The reason for this is that the provision of homologous PARPs would be particularly important for developing novel targets for drugs, and novel drugs, in order to improve

diagnosis and/or therapy of pathological states in which PARP, PARP homologs or substances derived therefrom are involved.

We have found that this object is achieved by providing PARP homologs having an amino acid sequence which has

a) a functional NAD^+ binding domain i.e. a PARP "signature" sequence having the characteristic GX_3GKG motif; and

b) especially in the N-terminal sequence region, i.e. in the region of the first 200, such as, for example, in the region of the first 100, N-terminal amino acids, no PARP zinc finger sequence motifs of the general formula

$$CX_2CX_mHX_2C$$

in which m is an integral value from 28 or 30, and the X radicals are, independently of one another, any amino acid;

and the functional equivalents thereof.

Since the PARP molecules according to the invention represent in particular functional homologs, they naturally also have a poly(ADP-ribose)-synthesizing activity. The NAD binding domain essentially corresponds to this activity and is localized to the C terminus.

Thus an essential characteristic of the PARPs according to the invention is the presence of a functional NAD^+ binding domain (PARP signature) which is located in the C-terminal region of the amino acid sequence (i.e. approximately in the region of the last 400, such as, for example, the last 350 or 300, C-terminal amino acids), in combination with an N-terminal sequence having no zinc finger motifs. Since the zinc finger motifs in known PARPs presumably contribute to recognition of the DNA breakages, it is to be assumed that the proteins according to the invention do not interact with DNA or do so in another way. It has been demonstrated by appropriate biochemical tests that the PARP2 according to the invention can be activated by 'activated DNA' (i.e. DNA after limited DNAaseI [sic] digestion). It can be concluded from this further that the PARP2 according to the invention have [sic] DNA binding properties. However, the mechanism of the DNA binding and enzyme activation differs between the PARPs according to the invention and PARP1. Its DNA binding and enzyme activation is, as mentioned, mediated by a characteristic zinc finger motif. No such motifs are present in the PARPs according to the invention. Presumably these properties are mediated by positively charged amino acids in the N-terminal region of the PARPs according to the invention. Since the 'activated DNA' (i.e. for example DNA after limited treatment

with DNAaseI [sic]) has a large number of defects (single strand breaks, single strand gaps, single-stranded overhangs, double strand breaks etc.), it is possible that although PARP1 and the PARPs according to the invention are activated by the same 'activated DNA', it is by a different subpopulation of defects (e.g. single strand breaks).

The functional NAD^+ binding domain (i.e. catalytic domain) binds the substrate for poly-ADP-ribose synthesis. Consistent with known PARPs, the sequence motif $GX^1X^2X^3GKG$, in which G is glycine, K is lysine, and X^1, X^2 and X^3 are, independently of one another, any amino acid, is present in particular. However, as shown, surprisingly, by comparison of the amino acid sequences of the NAD^+ binding domains of PARP molecules according to the invention with previously disclosed human PARP1, the sequences according to the invention differ markedly from the known sequence for the NAD^+ binding domain.

A group of PARP molecules which is preferred according to the invention preferably has the following general sequence motif in the catalytic domain in common:

$PX_n(S/T)GX_3GKGIYFA$, in particular
$(S/T)XGLR(I/V)XPX_n(S/T)GX_3GKGIYFA$, preferably
$LLWHG(S/T)X_7IL(S/T)XGLR(I/V)XPX_n(S/T)GX_3GKGIYFAX_3SKSAXY$

in which (S/T) describes the alternative occupation of this sequence position by S or T, (I/V) describes the alternative occupation of this sequence position by I or V, and n is an integral value from 1 to 5, and the X radicals are, independently of one another, any amino acid. The last motif is also referred to as the "PARP signature" motif.

The automodification domain is preferably likewise present in the PARPs according to the invention. It can be located, for example, in the region from about 100 to 200 amino acids in front of the N-terminal end of the NAD^+ binding domain.

PARP homologs according to the invention may additionally comprise, N-terminally of the NAD^+ binding domain (i.e. about 30 to about 80 amino acids closer to the N terminus), a leucine zipper-like sequence motif of the general formula

$$(L/V)X_6LX_6LX_6L$$

in which (L/V) represents the alternative occupation of this sequence position by L or V, and the X radicals are, independently of one another, any amino acid. The leucine zipper motifs observed according to the invention differ distinctly in

position from those described for PARP from Drosophila. Leucine zippers may lead to homodimers (two PARP molecules) or heterodimers (one PARP molecule with a binding partner differing therefrom).

The PARP homologs according to the invention preferably additionally comprise, N-terminally of the abovementioned leucine zipper-like sequence motif, i.e. about 10 to 250 amino acid residues closer to the N terminus, at least another one of the following part-sequence motifs:

$LX_9NX_2YX_2QLLX(D/E)X_bWGRVG$,	(motif 1)
$AX_3FXKX_4KTXNXWX_5FX_3PXK$,	(motif 2)
$QXL(I/L)X_2IX_9MX_{10}PLGKLX_3QIX_6L$,	(motif 3)
$FYTXIPHXFGX_3PP$,	(motif 4) and
$KX_3LX_2LXDIEXAX_2L$	(motif 5),

in which (D/E) describes the alternative occupation of this sequence position by D or E, (I/L) describes the alternative occupation of this sequence position by I or L, b is the integral value 10 or 11, and the X radicals are, independently of one another, any amino acid. It is most preferred for these motifs 1 to 5 all to be present in the stated sequence, with motif 1 being closest to the N terminus.

The abovementioned PARP signature motif is followed in the proteins according to the invention by at least another one of the following motifs:

$GX_3LXEVALG$	(motif 6)
$GX_2SX_4GX_3PX_aLXGX_2V$	(motif 7) and
$E(Y/F)X_2YX_3QX_4YLL$	(motif 8)

in which (Y/F) describes the alternative occupation of this sequence position by Y or F, a is equal to 7 to 9 and X is in each case any amino acid. It is most preferred for the three C-terminal motifs all to be present and in the stated sequence, with motif 8 being closest to the C terminus.

A preferred PARP structure according to the invention may be described schematically as follows:

Motifs 1 to 5/PARP signature/motifs 6 to 8 or
motifs 1 to 5/leucine zipper/PARP signature/motifs 6 to 8

it being possible for further amino acid residues, such as, for example, up to 40, to be arranged between the individual motifs and for further amino acid residues, such as, for example, up to 80, to be arranged at the N terminus and/or at the C terminus.

PARP homologs which are particularly preferred according to the invention are the proteins human PARP2, human PARP3, mouse PARP3 and the functional equivalents thereof. The proteins [sic] referred to as human PARP2 comprises 570 amino acids (cf. SEQ ID NO:2). The protein referred to as human PARP3 possibly exists in two forms. Type 1 comprises 533 amino acids (SEQ ID NO:4) and type 2 comprises 540 amino acids (SEQ ID NO:6). The forms may arise through different initiation of translation. The proteins [sic] referred to as mouse PARP3 exists in two forms which differ from one another by a deletion of 5 amino acids (15 bp). Type 1 comprises 533 amino acids (SEQ ID NO: 8) and type 2 comprises 528 amino acids (SEQ ID NO:10).

The invention further relates to the binding partners for the PARP homologs according to the invention. These binding partners are preferably selected from

a) antibodies and fragments such as, for example, Fv, Fab, $(Fab)'_2$ [sic], thereof
b) protein-like compounds which interact, for example via the above leucine zipper region or another sequence section, with PARP, and
c) low molecular weight effectors which modulate a biological PARP function such as, for example, the catalytic PARP activity, i.e. NAD^+-consuming ADP ribosylation, or the binding to an activator protein or to DNA.

The invention further relates to nucleic acids comprising

a) a nucleotide sequence coding for at least one PARP homolog according to the invention, or the complementary nucleotide sequence thereof;
b) a nucleotide sequence which hybridizes with a sequence as specified in a), preferably under stringent conditions; or
c) nucleotide sequences which are derived from the nucleotide sequences defined in a) and b) through the degeneracy of the genetic code.

Nucleic acids which are suitable according to the invention comprise in particular at least one of the part-sequences which code for the abovementioned amino acid sequence motifs.

Nucleic acids which are preferred according to the invention comprise nucleotide sequences as shown in SEQ ID NO: 1 and 3, and, in particular, part-sequences thereof which are characteristic of PARP homologs according to the invention, such as, for example, nucleotide sequences comprising

a) nucleotides +3 to +1715 shown in SEQ ID NO:1;
b) nucleotides +242 to +1843 shown in SEQ ID NO:3;
c) nucleotides +221 to +1843 shown in SEQ ID NO:5;
d) nucleotides +112 to +1710 shown in SEQ ID NO:7; or
e) nucleotides +1 to +1584 shown in SEQ ID NO:9

or part-sequences of a), b), c), d) and e) which code for the abovementioned characteristic amino acid sequence motifs of the PARP homologs according to the invention.

The invention further relates to expression cassettes which comprise at least one of the above-described nucleotide sequences according to the invention under the genetic control of regulatory nucleotide sequences. These can be used to prepare recombinant vectors according to the invention, such as, for example, viral vectors or plasmids, which comprise at least one expression cassette according to the invention.

Recombinant microorganisms according to the invention are transformed with at least one of the abovementioned vectors.

The invention also relates to transgenic mammals transfected with a vector according to the invention.

The invention further relates to an in vitro detection method, which can be carried out homogeneously or heterogeneously, for PARP inhibitors, which comprises

a) incubating an unsupported or supported polyADP-ribosylatable target with a reaction mixture comprising
   a1) a PARP homolog according to the invention;
   a2) a PARP activator; and
   a3) a PARP inhibitor or an analyte in which at least one PARP inhibitor is suspected;
b) carrying out the polyADP ribosylation reaction; and
c) determining the polyADP ribosylation of the target qualitatively or quantitatively.

The detection method is preferably carried out by preincubating the PARP homolog with the PARP activator and the PARP inhibitor or an analyte in which at least one PARP inhibitor is suspected,

for example for about 1-30 minutes, before carrying out the poly-ADP ribosylation reaction.

After activation by DNA with single strand breaks (referred to as "activated DNA" according to the invention), PARP polyADP ribosylates a large number of nuclear proteins in the presence of NAD. These proteins include, on the one hand, PARP itself, but also histones etc.

The polyADP-ribosylatable target preferably used in the detection method is a histone protein in its native form or a polyADP-ribosylatable equivalent derived therefrom. A histone preparation supplied by Sigma (SIGMA, catalogue No. H-7755; histone type II-as [sic] from calf thymus, Luck, J. M., et al., J. Biol. Chem., 233, 1407 (1958), Satake K., et al., J. Biol. Chem, 235, 2801 (1960)) was used by way of example. It is possible in principle to use all types of proteins or parts thereof amenable to poly-ADP-ribosylation by PARP. These are preferably nuclear proteins, e.g. histones, DNA polymerase, telomerase or PARP itself. Synthetic peptides derived from the corresponding proteins can also act as target.

In the ELISA assay according to the invention it is possible to use amounts of histones in the range from about 0.1 µg/well to about 100 µg/well, preferably about 1 µg/well to about 10 µg/well. The amounts of the PARP enzyme are in a range from about 0.2 pmol/well to about 2 nmol/well, preferably from about 2 pmol/well to about 200 pmol/well, the reaction mixture comprising in each case 100 µg/well. Reductions to smaller wells and correspondingly smaller reaction volumes are possible.

In the HTRF assay according to the invention, identical amounts of PARP are employed, and the amount of histone or modified histones is in the range from about 2 ng/well to about 25 µg/well, preferably about 25 ng/well to about 2.5 µg/well, the reaction mixture comprising in each case 50 µl/well. Reductions to smaller wells and correspondingly smaller reaction volumes are possible.

The PARP activator used according to the invention is preferably activated DNA.

Various types of damaged DNA can function as activator. DNA damage can be produced by digestion with DNAases [sic] or other DNA-modifying enzymes (e.g. restriction endonucleases), by irradiation or other physical methods or chemical treatment of the DNA. It is further possible to simulate the DNA damage situation in a targeted manner using synthetic oligonucleotides. In the assays

indicated by way of example, activated DNA from calf thymus was employed (Sigma, product No. D4522; CAS: 91080-16-9, prepared by the method of Aposhian and Kornberg using calf thymus DNA (SIGMA D-1501) and deoxyribonuclease type I (D-4263). Aposhian H. V. and Kornberg A., J. Biol. Chem., 237, 519 (1962)). The activated DNA was used in a concentration range from 0.1 to 1000 µg/ml, preferably from 1 to 100 µg/ml, in the reaction step.

The polyADP ribosylation reaction is started in the method according to the invention by adding NAD^+. The NAD concentrations were in a range from about 0.1 µM to about 10 mM, preferably in a range from about 10 µM to about 1 mM.

In the variant of the above method which can be carried out heterogeneously, the polyADP ribosylation of the supported target is determined using anti-poly(ADP-ribose) antibodies. To do this, the reaction mixture is separated from the supported target, washed and incubated with the antibody. This antibody can itself be labeled. However, it is preferable to use for detecting bound anti-poly(ADP-ribose) antibody a labeled secondary antibody or a corresponding labeled antibody fragment. Suitable labels are, for example, radiolabeling, chromophore- or fluorophore-labeling, biotinylation, chemiluminescence labeling, labeling with paramagnetic material or, in particular, enzyme labels, e.g. with horseradish peroxidase. Appropriate detection techniques are generally known to the skilled worker.

In the variant of the above process which can be carried out homogeneously, the unsupported target is labeled with an acceptor fluorophore. The target preferably used in this case is biotinylated histone, the acceptor fluorophore being coupled via avidin or streptavidin to the biotin groups of the histone. Particularly suitable as acceptor fluorophore are phycobiliproteins (e.g. phycocyanins, phycoerythrins), e.g. R-phycocyanin (R-PC), allophycocyanin (APC), R-phycoerythrin (R-PE), C-phycocyanin (C-PC), B-phycoerythrin (B-PE) or their combinations with one another or with fluorescent dyes such as Cy5, Cy7 or Texas Red (Tandem system) (Thammapalerd, N. et al., Southeast Asian Journal of Tropical Medicine & Public Health, 27(2): 297-303 (1996); Kronick, M. N. et al., Clinical Chemistry, 29(9), 1582-1586 (1986); Hicks, J. M., Human Pathology, 15(2), 112-116 (1984)). The dye XL665 used in the examples is a crosslinked allophycocyanin (Glazer, A. N., Rev. Microbiol., 36, 173-198 (1982); Kronick, M. N., J. Imm. Meth., 92, 1-13 (1986); MacColl, R. et al., Phycobiliproteins, CRC Press, Inc., Boca Raton, Florida (1987); MacColl, R. et al., Arch. Biochem. Biophys., 208(1), 42-48 (1981)).

It is additionally preferred in the homogeneous method to determine the polyADP ribosylation of the unsupported target using anti-poly(ADP-ribose) antibody which is labeled with a donor fluorophore which is able to transfer energy to the acceptor fluorophore when donor and acceptor are close in space owing to binding of the labeled antibody to the polyADP-ribosylated histone. A europium cryptate is preferably used as donor flurophore for the anti-poly(ADP-ribose) antibody.

Besides the europium cryptate used, other compounds are also possible as potential donor molecules. This may entail, on the one hand, modification of the cryptate cage. Replacement of the europium by other rare earth metals such as terbium are [sic] also conceivable. It is crucial that the fluorescence has a long duration to guarantee the time delay (Lopez, E. et al., Clin. Chem. 39/2, 196-201 (1993); US Patent 5,534,622).

The detection methods described above are based on the principle that there is a correlation between the PARP activity and the amount of ADP-ribose polymers formed on the histones. The assay described herein makes it possible to quantify the ADP-ribose polymers using specific antibodies in the form of an ELISA and an HTRF (homogenous time-resolved fluorescence) assay. Specific embodiments of these two assays are described in detail in the following examples.

The developed HTRF (homogeneous time-resolved fluorescence) assay system measures the formation of poly(ADP-ribose) on histones using specific antibodies. In contrast to the ELISA, this assay is carried out in homogeneous phase without separation and washing steps. This makes a higher sample throughput and a smaller susceptibility to errors possible. HTRF is based on the fluorescence resonance energy transfer (FRET) between two fluorophores. In a FRET assay, an excited donor fluorophore can transfer its energy to an acceptor fluorophore when the two are close to one another in space. In HTRF technology, the donor fluorophore is a europium cryptate [(Eu)K] and the acceptor is XL665, a stabilized allophycocyanin. The europium cryptate is based on studies by Jean Marie Lehn (Strasbourg) (Lopez, E. et al., Clin. Chem. 39/2, 196-201 (1993); US Patent 5,534,622).

In a homogeneous assay, all the components are also present during the measurement. Whereas this has advantages for carrying out the assay (rapidity, complexity), it is necessary to preclude interference by assay components (inherent fluorescence, quenching by dyes etc.). HTRF precludes such interference by time-delayed measurement at two wavelengths (665 nm, 620 nm). The HTRF fluor-

escence [sic] has a very long decay time and time-delayed measurement is therefore possible. There is no longer any interference from short-lived background fluorescence (e.g. from assay components or inhibitors of the substance bank). In addition, measurement is always carried out at two wavelengths in order to compensate for quench effects of colored substances. HTRF assays can be carried out, for example, in 96- or 384-well microtiter plate format and are evaluated using a discovery HTRF microplate analyzer (Packard Instruments).

Also provided according to the invention are the following in vitro screening methods for binding partners for PARP, in particular for a PARP homolog according to the invention.

A first variant is carried out by
a1) immobilizing at least one PARP homolog on a support;
b1) contacting the immobilized PARP homolog with an analyte in which at least one binding partner is suspected; and
c1) determining, where appropriate after an incubation period, analyte constituents bound to the immobilized PARP homolog.

A second variant entails
a2) immobilizing on a support an analyte which comprises at least one possible binding partner for the PARP homolog;
b2) contacting the immobilized analyte with at least one PARP homolog for which a binding partner is sought; and
c3) examining the immobilized analyte, where appropriate after an incubation period, for binding of the PARP homolog.

The invention also relates to a method for the qualitative or quantitative determination of a nucleic acid encoding a PARP homolog, which comprises
a) incubating a biological sample with a defined amount of an exogenous nucleic acid according to the invention (e.g. with a length of about 20 to 500 bases or longer), hybridizing, preferably under stringent conditions, determining the hybridizing nucleic acids and, where appropriate, comparing with a standard; or

b) incubating a biological sample with a defined amount of oligonucleotide primer pairs with specificity for a PARP homolog-encoding nucleic acid, amplifying the nucleic acid, determining the amplification product and, where appropriate, comparing with a standard.

The invention further relates to a method for the qualitative or quantitative determination of a PARP homolog according to the invention, which comprises

a) incubating a biological sample with at least one binding partner specific for a PARP homolog,
b) detecting the binding partner/PARP complex and, where appropriate,
c) comparing the result with a standard.

The binding partner in this case is preferably an anti-PARP antibody or a binding fragment thereof, which carries a detectable label where appropriate.

The determination methods according to the invention for PARP, in particular for PARP homologs and for the coding nucleic acid sequences thereof, are suitable and advantageous for diagnosing sepsis- or ischemia-related tissue damage, in particular strokes, myocardial infarcts or septic shock.

The invention further comprises a method for determining the efficacy of PARP effectors, which comprises

a) incubating a PARP homolog according to the invention with an analyte which comprises an effector of a physiological or pathological PARP activity; removing the effector again where appropriate; and
b) determining the activity of the PARP homolog, where appropriate after adding substrates or cosubstrates.

The invention further relates to gene therapy compositions which comprise in a vehicle acceptable for gene therapy a nucleic acid construct which

a) comprises an antisense nucleic acid against a coding nucleic acid according to the invention; or
b) a ribozyme against a noncoding nucleic acid according to the invention; or
c) codes for a specific PARP inhibitor.

The invention further relates to pharmaceutical compositions comprising, in a pharmaceutically acceptable vehicle, at least one PARP protein according to the invention, at least one PARP binding partner according to the invention or at least one coding nucleotide sequence according to the invention.

Finally, the invention relates to the use of binding partners of a PARP homolog for the diagnosis or therapy of pathological states in the development and/or progress of which at least one PARP protein, in particular a PARP homolog according to the invention

invention, or a polypeptide derived therefrom, are [sic] involved. The binding partner used can be, for example, a low molecular weight binding partner whose molecular weight can be, for example, less than about 2000 dalton or less than about 1000 dalton.

The invention additionally relates to the use of PARP binding partners for the diagnosis or therapy of pathological states mediated by an energy deficit. An energy deficit for the purpose of the present invention is, in particular, a cellular energy deficit which is to be observed in the unwell patient systemically or in individual body regions, ogans [sic] or organ regions, or tissues or tissue regions. This is characterized by an NAD and/or ATP depletion going beyond the physiological range of variation of the NAD and/or ATP level and mediated preferably by a protein with PARP activity, in particular a PARP homolog according to the invention, or a polypeptide derived therefrom.

The invention particularly relates to the use of a PARP binding partner as defined above for the diagnosis or therapy (acute or prophylactic) of pathological states mediated by energy deficits and selected from neurodegenerative disorders, or tissue damage caused by sepsis or ischemia, in particular of neurotoxic disturbances, strokes, myocardial infarcts, damage which [sic] during or after infarct lysis with drugs (B. [sic] with TPA, Reteplase or mechanically with laser or Rotablator) and of microinfarcts during and after heart valve replacement, aneurysm resections and heart transplants, trauma to the head and spinal cord, infarcts of the kidney (acute kidney failure, acute renal insufficiency or damage during and after kidney transplant), infarcts of the liver (liver failure, damage during or after a liver transplant), peripheral neuropathies, AIDS dementia, septic shock, diabetes, trauma (craniocerebral trauma), massive bleeding, subarachnoid hemorrhages, Alzheimer's disease, multi-infarct dementia, Huntington's disease, epilepsy, Parkinson's disease, amyotrophic lateral sclerosis, kidney failure, also in the chemotherapy of tumors and prevention of metastasis and for the treatment of inflammations and rheumatic disorders, e.g. of rheumatoid arthritis; further for the treatment of revascularization of critically narrowed coronary arteries and critically narrowed peripheral arteries, e.g. leg arteries.

Nonlimiting examples of tumors are leukemia, glioblastomas, lymphomas, melanomas, carcinomas of the breast and cervix etc.

The present invention will now be described in more detail with reference to the appended figures. These show:

In Figure 1 a sequence alignment of human PARP (human PARP1) and two PARPs preferred according to the invention (human PARP2, human PARP3, murine PARP3). Sequence agreements between human PARP1 and human PARP2, human PARP3 or murine PARP3 are depicted within frames. The majority sequence is indicated over the alignment. The zinc finger motifs of human PARP1 are located in the sequence sections corresponding to amnio [sic] acid residues 21 to 56 and 125 to 162;

In Figure 2 Northern blots with various human tissues to illustrate the tissue distribution of PARP2 and PARP3 molecules according to the invention. Lane 1: brain; lane 2: heart; lane 3: skeletal muscle; lane 4: large bowel; lane 5: thymus; lane 6: spleen; lane 7: kidney; lane 8: liver; lane 9: small bowel; lane 10: placenta; lane 11: lung; lane 12: peripheral blood leukocytes; the respective position of the size standard (kb) is indicated.

In Figure 3 a Northern blot with further various human tissues to illustrate the tissue distribution of the PARP3 molecule according to the invention. Lane 1: heart; lane 2: brain; lane 3: placenta; lane 4: lung; lane 5: liver; lane 6: skeletal muscle; lane 7: kidney; lane 8: pancreas; the respective position of the size standard (kD [sic]) is indicated.

In Figure 4 a Western blot with various human tissues to illustrate the tissue distribution of the PARP3 molecule according to the invention at the protein level. Lane 1: heart; lane 2: lung; lane 3: liver; lane 4: spleen; lane 5: kidney; lane 6: large bowels; lane 7: muscle; lane 8: brain; the respective position of the size standard (kD) is indicated.

In Figure 5 a Western blot with various human tissues to illustrate the tissue distribution of the PARP3 molecule according to the invention. Lane 1: frontal cortex; lane 2: posterior cortex; lane 3: cerebellum; lane 4: hippocampus; lane 5: olfactory bulb; lane 6: striatum; lane 7: thalamus; lane 8: midbrain; lane 9: entorhinal cortex; lane 10: pons; lane 11: medulla; lane 12: spinal cord.

In Figure 6 a diagrammatic representation of the PARP assay (ELISA)

In Figure 7 a diagrammatic representation of the PARP assay (HTRF)

Further preferred embodiments of the invention are described in the following sections.

PARP homologs and functional equivalents

Unless stated otherwise, for the purposes of the present description amino acid sequences are indicated starting with the N terminus. If the one-letter code is used for amino acids, then G is glycine, A is alanine, V is valine, L is leucine, I is isoleucine, S is serine, T is threonine, D is aspartic acid, N is asparagine, E is glutamic acid, Q is glutamine, W is tryptophan, H is histidine, R is arginine, P is proline, K is lysine, Y is tyrosine, F is phenylalanine, C is cysteine and M is methionine.

The present invention is not confined to the PARP homologs specifically described above. On the contrary, those homologs which are functional equivalents thereof are also embraced. Functional equivalents comprise both natural, such as, for example, species-specific or organ-specific, and artificially produced variants of the proteins specifically described herein. Functional equivalents according to the invention differ by addition, substitution, inversion, insertion and/or deletion of one or more amino acid residues of human PARP2 (SEQ ID NO:2), human PARP3 (SEQ ID NO: 4 and 6) and mouse PARP3 (SEQ ID:8 and 10), there being at least retention of the NAD-binding function of the protein mediated by a functional catalytic C-terminal domain. Likewise, the poly(ADP-ribose)-producing catalytic activity should preferably be retained. Functional equivalents also comprise where appropriate those variants in which the region similar to the leucine zipper is essentially retained.

It is moreover possible, for example, starting from the sequence for human PARP2 or human PARP3 to replace certain amino acids by those with similar physicochemical properties (bulk, basicity, hydrophobicity, etc.). It is possible, for example, for arginine residues to be replaced by lysine residues, valine residues by isoleucine residues or aspartic acid residues by glutamic acid residues. However, it is also possible for one or more amino acids to be exchanged in sequence, added or deleted, or several of these measures can be combined together. The proteins which have been modified in this way from the human PARP2 or human PARP3 sequence have at least 60%, preferably at least 75%, very particularly preferably at least 85%, homology with the starting

sequence, calculated using the algorithm of Pearson and Lipman, Proc. Natl. Acad. Sci (USA) 85(8), 1988, 2444-2448.

The following homologies have been determined at the amino acid level and DNA level between human PARP1, 2 and 3 (FastA program, Pearson and Lipman, loc. cit.):
Amino acid homologies:

	Percent identity	Percent identity in PARP signature
PARP1/PARP2	41.97% (517)	86% (50)
PARP1/PARP3	33.81% (565)	53.1% (49)
PARP2/PARP3	35.20% (537)	53.1% (49)

Numbers in parentheses indicate the number of overlapping amino acids.

DNA Homologies:

	Percent identity in the ORF	Percent identity in PARP signature
PARP1/PARP2	60.81% (467)	77.85% (149)
PARP1/PARP3	58.81% (420)	59.02% (61)
PARP2/PARP3	60.22% (269)	86.36% (22)

Numbers in parentheses indicate the number of overlapping nucleotides.

The polypeptides according to the invention can be classified as homologous poly(ADP-ribose) polymerases on the basis of the great similarity in the region of the catalytic domain.

It is also essential to the invention that the novel PARP homologs do not have conventional zinc finger motifs. This means that these enzymes is [sic] not necessarily involved in DNA repair or is [sic] so in a way which prevents from [sic] PARP1, but are still able to carry out their pathological mechanism (NAD^+ consumption and thus energy consumption due to ATP consumption). The strong protein expression, particularly of PARP3, observable in the Western blot suggests a significant role in the NAD

consumption. This is particularly important for drug development. Potential novel inhibitors of the polymerases according to the invention can thus inhibit the pathological functions without having adverse effects on the desired physiological properties. This was impossible with inhibitors against the PARPs known to date since there was always also inhibition of the DNA repair function. The potentially mutagenic effect of known PARP inhibitors is thus easy to understand. It is also conceivable to design PARP inhibitors so that they efficiently inhibit all PARP homologs with high affinity. In this case, a potentiated effect is conceivable where appropriate.

The PARP homolog which is preferred according to the invention and is shown in SEQ ID NO:2 (human PARP2) can advantageously be isolated from human brain, heart, skeletal muscle, kidney and liver. The expression of human PARP2 in other tissues or organs is distinctly weaker.

The PARP homolog which is preferred according to the invention and is shown in SEQ ID NO: 4 and 6 (human PARP3) can advantageously be isolated from human brain (in this case very specifically from the hippocampus), heart, skeletal muscle, liver or kidney. The expression of human PARP3 in other tissues or organs, such as muscle or liver, is distinctly weaker.

The skilled worker familiar with protein isolation will make use of the combination of preparative methodologies which is most suitable in each case for isolating natural PARPs according to the invention from tissues or recombinantly prepared PARPs according to the invention from cell cultures. Suitable standard preparative methods are described, for example, in Cooper, T.G., Biochemische Arbeitsmethoden, published by Walter de Gruyter, Berlin, New York or in Scopes, R. Protein Purification, Springer Verlag, New York, Heidelberg, Berlin.

The invention additionally relates to PARP2 and PARP3 homologs which, although they can be isolated from other eukaryotic species, i.e. invertebrates or vertebrates, especially other mammals such as, for example, mice, rats, cats, dogs, pigs, sheep, cattle, horses or monkeys, or from other organs such as, for example the myocardium, have the essential structural and functional properties predetermined by the PARPs according to the invention.

In particular, the human PARP2 which can be isolated from human brain, and its functional equivalents, are preferred agents for developing inhibitors of stroke. This is because it can be

assumed that drug development based on PARP2 as indicator makes it possible to develop inhibitors which are optimized for use in the human brain. However, it cannot be ruled out that inhibitors developed on the basis of PARP2 can also be employed for treating PARP-mediated pathological states in other organs too. On the basis of the tissue distribution of the proteins according to the invention, indications of particular interest are those derived from ischemic states of appropriate organs (ischemia of the brain (stroke), of the heart (myocardial infarct), damage which [sic] during or after infarct lysis (e.g. with TPA, Reteplase or mechanically with laser or Rotoblator) and from [sic] microinfarcts during and after heart valve replacement, aneurysm resections and heart transplants, of the kidney (acute kidney failure, acute renal insufficiency or damage during and after a kidney transplant), damage to the liver or the skeletal muscle). Also conceivable are the treatment and prophylaxis of neurodegenerative disorders occurring after ischemia, trauma, (craniocerebral trauma), massive bleeding, subarachnoid hemorrhages and stroke, and of neurodegenerative disorders such as multi-infarct dementia, Alzheimer's disease, Huntington's disease and epilepsies, especially of generalized epileptic seizures such as petit mal and tonoclonic seizures and partial epileptic seizures, such as temporal lope [sic], and complex partial seizures. Said proteins may also be relevant for the treatment of revascularization of critical narrowed coronary arteries and critically narrowed peripheral arteries, e.g. leg arteries. Said proteins may additionally play a part in the chemotherapy of tumors and in the prevention of metastasis, and in the treatment of inflammations and rheumatic disorders, e.g. of rheumatoid arthritis. Further pathological states of these and other organs are conceivable.

PARP2 and 3 are also, similar to PARP1, activated by damaged DNA, although by a presumably different mechanism. Significance in DNA repair is conceivable. Blockade of the PARPs according to the invention would also be beneficial in indications such as cancer (e.g. in the radiosensitization of tumor patients).

Another essential biological property of PARPs according to the invention and their functional equivalents is to be seen in their ability to bind an interacting partner. Human PARP2 and 3 differ from previously disclosed PARPs from higher eukaryotes such as, in particular, mammals by having potential so-called leucine zipper motifs. This is a typical motif for protein-protein interactions. It is possible that these motifs permit modulation of PARP activity by an interacting partner. This additional structural element thus also provides a possible starting point

for development of PARP effectors such as, for example, inhibitors.

The invention thus further relates to proteins which interact with PARP2 and/or 3, preferably those which bring about their activation or inactivation.

The invention further relates to proteins which still have the abovementioned ligand-binding activity and which can be prepared starting from the specifically disclosed amino acid sequences by targeted modifications.

It is possible, starting from the peptide sequence of the proteins according to the invention, to generate synthetic peptides which are employed, singly or in combination, as antigens for producing polyclonal or monoclonal antibodies. It is also possible to employ the PARP protein or fragments thereof for generating antibodies. The invention thus also relates to peptide fragments of PARP proteins according to the invention which comprise characteristic part-sequences, in particular those oligo- or polypeptides which comprises [sic] at least one of the abovementioned sequence motifs. Fragments of this type can be obtained, for example, by proteolytic digestion of PARP proteins or by chemical synthesis of peptides.

Novel specific PARP binding partners

Active and selective inhibitors against the proteins according to the invention were developed using the specific assay systems described above for binding partners for PARP1, PARP2 and PARP3.

Inhibitors provided according to the invention have a strong inhibitory activity on PARP2. The K_i values may in this case be less than about 100 nM, such as less than about 700 nM, less than about 100 nM and less than about 30 nM, e.g. about 1 to 20 nM.

Inhibitors preferred according to the invention have a surprisingly marked selectivity for PARP2. This is shown by the K_i(PARP1) : K_i(PARP2) ratio for inhibitors according to the invention which is, for example, greater than 5, preferably greater than 10 and, in particular, greater than 20 and is, for example, in the range of about 30-100, e.g. about 40-80. Another group of inhibitors was developed so that they inhibit PARP1 and PARP2 simultaneously.

An example which should be mentioned is 2(4(2-(N,N-diethylamino)eth-1-yloxy)phenyl)benzimidazole-4-carboxamide [sic]. This compound showed selectivity for PARP2 (K_i=7nM) relative to PARP1 (K_i=200nM).

Nucleic acids coding for PARP homologs:

Unless stated otherwise, nucleotide sequences are indicated in the present description from the 5′ to the 3′ direction.

The invention further relates to nucleic acid sequences which code for the abovementioned proteins, in particular for those having the amino acid sequence depicted in SEQ ID NO: 2, 4, 6, 8 and 10, but without being restricted thereto. Nucleic acid sequences which can be used according to the invention also comprise allelic variants which, as described above for the amino acid sequences, are obtainable by deletion, inversion, insertion, addition and/or substitution of nucleotides, preferably of nucleotides shown in SEQ ID NO: 1, 3, 7 and 9, but with essential retention of the biological properties and the biological activity of the corresponding gene product. Nucleotide sequences which can be used are obtained, for example, by silent (without alteration of the amino acid sequence) or conservative (exchange of amino acids of the same size, charge, polarity or solubility) nucleotide substitutions.

Nucleic acid sequences according to the invention also embrace functional equivalents of the genes, such as eukaryotic homologs for example from invertebrates such as Caenorhabditis or Drosophila, or vertebrates, preferably from the mammals described above. Preferred genes are those from vertebrates which code for a gene product which has the properties essential to the invention as described above.

The nucleic acids according to the invention can be obtained in a conventional way by various routes:

For example, a genomic or a cDNA library can be screened for DNA which codes for a PARP molecule or a part thereof. For example, a DNA library obtained from human brain, heart or kidney can be screened with a suitable probe such as, for example, a labeled single-stranded DNA fragment which corresponds to a part-sequence of suitable length selected from SEQ ID NO: 1 or 3, or sequence complementary thereto. For this purpose, it is possible, for example, for the DNA fragments of the library which have been transferred into a suitable cloning vector to be, after transformation into a bacterium, plated out on agar plates. The

clones can then be transferred to nitrocellulose filters and, after denaturation of the DNA, hybridized with the labeled probe. Positive clones are then isolated and characterized.

The DNA coding for PARP homologs according to the invention or partial fragments can also be synthesized chemically starting from the sequence information contained in the present application. For example, it is possible for this purpose for oligonucleotides with a length of about 100 bases to be synthesized and sequentially ligated in a manner known per se by, for example, providing suitable terminal restriction cleavage sites.

The nucleotide sequences according to the invention can also be prepared with the aid of the polymerase chain reaction (PCR). For this, a target DNA such as, for example, DNA from a suitable full-length clone is hybridized with a pair of synthetic oligonucleotide primers which have a length of about 15 bases and which bind to opposite ends of the target DNA. The sequence section lying between them is then filled in with DNA polymerase. Repetition of this cycle many times allows the target DNA to be amplified (cf. White et al.(1989), Trends Genet. 5, 185).

The nucleic acid sequences according to the invention are also to be understood to include truncated sequences, single-stranded DNA or RNA of the coding and noncoding, complementary DNA sequence, mRNA sequences and cDNAs derived therefrom.

The invention further embraces nucleotide sequences hybridizing with the above sequences under stringent conditions. Stringent hybridization conditions for the purpose of the present invention exist when the hybridizing sequences have a homology of about 70 to 100%, such as, for example about 80 to 100% or 90 to 100% (preferably in an amino acid section of at least about 40, such as, for example, about 50, 100, 150, 200, 400 or 500 amino acids).

Stringent conditions for the screening of DNA, in particular cDNA banks, exist, for example, when the hybridization mixture is washed with 0.1X SSC buffer (20X SSC buffer = 3M NaCl, 0.3M sodium citrate, pH 7.0) and 0.1% SDS at a temperature of about 60°C.

Northern blot analyses are analyses are washed under stringent conditions with 0.1X SSC, 0,1% SDS at a temperature of about 65°C, for example.

Nucleic acid derivatives and expression constructs:

The nucleic acid sequences are also to be understood to include derivatives such as, for example, promoter variants or alternative splicing variants. The promoters operatively linked in front of the nucleotide sequences according to the invention may moreover be modified by nucleotide addition(s) or substitution(s), inversion(s), insertion(s) and/or deletion(s), but without impairing the functionality or activity of the promoters. The promoters can also have their activity increased by modifying their sequence, or be completely replaced by more effective promoters even from heterologous organisms. The promoter variants described above are used to prepare expression cassettes according to the invention.

Specific examples of human PARP2 splicing variants which may be mentioned are:

Variant human PARP2a: Deletion of base pairs 766 to 904 (cf. SEQ ID NO:1). This leads to a frame shift with a new stop codon ("TAA" corresonding to nucleotides 922 to 924 in SEQ ID NO:1).
Variant human PARP2b: Insertion of
5'- gta tgc cag gaa ggt cat ggg cca gca aaa ggg tct ctg -3'
after nucleotide 204 (SEQ ID NO:1). This extends the amino acid sequence by the insertion: GMPGRSWASKRVS

Nucleic acid derivatives also mean variants whose nucleotide sequence [sic] in the region from -1 to -1000 in front of the start codon have been modified so that gene expression and/or protein expression is increased.

Besides the nucleotide sequence described above, the nucleic acid constructs which can be used according to the invention comprise in functional, operative linkage one or more other regulatory sequences, such as promoters, amplification signals, enhancers, polyadenylation sequences, origins of replication, reporter genes, selectable marker genes and the like. This linkage may, depending on the desired use, lead to an increase or decrease in gene expression.

In addition to the novel regulatory sequences, it is possible for the natural regulatory sequence still to be present in front of the actual structural genes. This natural regulation can, where appropriate, be switched off by genetic modification, and the expression of the genes increased or decreased. However, the gene construct may also have a simpler structure, that is to say no additional regulatory signals are inserted in front of the

structural genes, and the natural promoter with its regulation is not deleted. Instead, the natural regulatory sequence is mutated in such a way that regulation no longer takes place, and gene expression is enhanced or diminished. It is also possible to insert additional advantageous regulatory elements at the 3' end of the nucleic acid sequences. The nucleic acid sequences can be present in one or more copies in the gene construct.

Advantageous regulatory sequences for the expression method according to the invention are, for example, present in promoters such as cos, tac, trp, tet, trp-tet, lpp, lac, lpp-lac, lacIq, T7, T5, T3, gal, trc, ara, SP6, l-PR or the l-PL promoter, which are advantageously used in Gram-negative bacteria. Other advantageous regulatory sequences are present, for example, in the Gram-positive promoters amy and SPO2, in the yeast promoters ADC1, MFa, AC, P-60, CYC1, GAPDH or in the plant promoters CaMV/35S, SSU, OCS, lib4, usp, STLS1, B33, nos or in the ubiquitin or phaseolin promoter.

It is possible in principle to use all natural promoters with their regulatory sequences. It is also possible and advantageous to use synthetic promoters.

Said regulatory sequences are intended to make specific expression of the nucleic acid sequences and of [sic] protein expression possible. This may mean, for example, depending on the host organism that the gene is expressed or overexpressed only after induction, or that it is immediately expressed and/or overexpressed.

The regulatory sequences or factors may moreover preferably have a positive influence on, and thus increase or decrease, the expression. Thus, enhancement of the regulatory elements may advantageously take place at the level of transcription by using strong transcription signals such as promoters and/or enhancers. However, it is also possible to enhance translation by, for example, improving the stability of the mRNA.

Enhancers mean, for example, DNA seqences which bring about increased expression via an improved interaction between RNA polymerase and DNA.

The recombinant nucleic acid construct or gene construct is, for expression in a suitable host organism, advantageously inserted into a host-specific vector which makes optimal expression of the genes in the host possible. Vectors are well known to the skilled worker and are to be found, for example, in "Cloning Vectors"

(Pouwels P. H. et al., Ed., Elsevier, Amsterdam-New York-Oxford, 1985). Apart from plasmids, vectors also mean all other vectors known to the skilled worker, such as, for example, phages, viruses, such as SV40, CMV, baculovirus and adenovirus, transposons, IS elements, phasmids, cosmids, and linear or circular DNA. These vectors may undergo autonomous replication in the host organism or chromosomal replication.

Expression of the constructs:

The recombinant constructs according to the invention described above are advantageously introduced into a suitable host system and are expressed. Cloning and transfection methods familiar to the skilled worker are preferably used in order to bring about expression of said nucleic acids in the particular expression system. Suitable systems are described, for example, in Current Protocols in Molecular Biology, F. Ausubel et al., ed., Wiley Interscience, New York 1997.

Suitable host organisms are in principle all organisms which make it possible to express the nucleic acids according to the invention, their allelic variants, their functional equivalents or derivatives or the recombinant nucleic acid construct. Host organisms mean, for example, bacteria, fungi, yeasts, plant or animal cells. Preferred organisms are bacteria such as those of the genera Escherichia, such as, for example, Escherichia coli, Streptomyces, Bacillus or Pseudomonas, eukaryotic microorganisms such as Saccharomyces cerevisiae, Aspergillus, higher eukaryotic cells from animals or plants, for example Sf9 or CHO cells.

The gene product can also, if required, be expressed in transgenic organisms such as transgenic animals such as, in particular, mice, sheep, or transgenic plants. The transgenic organisms may also be so-called knock-out animals or plants in which the corresponding endogenous gene has been switched off, such as, for example, by mutation or partial or complete deletion.

The combination of the host organisms and the vectors appropriate for the organisms, such as plasmids, viruses or phages, such as, for example, plasmids with the RNA polymerase/promoter system, phages λ, μ or other temperate phages or transposons and/or other advantageous regulatory sequences form [sic] an expression system. The term expression systems preferably means, for example, a combination of mammalian cells such as CHO cells, and

vectors, such as pcDNA3neo vector, which are suitable for mammalian cells.

As described above, the gene product can also be expressed advantageously in transgenic animals, e.g. mice, sheep, or transgenic plants. It is likewise possible to program cell-free translation systems with the RNA derived from the nucleic acid.

The gene product can also be expressed in the form of therapeutically or diagnostically suitable fragments. To isolate the recombinant protein it is possible and advantageous to use vector systems or oligonucleotides which extend the cDNA by certain nucleotide sequences and thus code for modified polypeptides which serve to simplify purification. Suitable modifications of this type are, for example, so-called tags which act as anchors, such as, for example, the modification known as the hexa-histidine anchor, or epitopes which can be recognized as antigens by antibodies (described, for example, in Harlow, E. and Lane, D., 1988, Antibodies: A Laboratory Manual. Cold Spring Harbor (N.Y.) Press). These anchors can be used to attach the proteins to a solid support such as, for example, a polymer matrix, which can, for example, be packed into a chromatography column, or to a microtiter plate or to another support.

These anchors can also at the same time be used to recognize the proteins. It is also possible to use for recognition of the proteins conventional markers such as fluorescent dyes, enzyme markers which form a detectable reaction product after reaction with a substrate, or radioactive markers, alone or in combination with the anchors for derivatizing the proteins.

Production of antibodies:

Anti-PARP2 antibodies are produced in a manner familiar to the skilled worker. Antibodies mean both polyclonal, monoclonal, human or humanized antibodies or fragments thereof, single chain antibodies or else synthetic antibodies, likewise antibody fragments such as Fv, Fab and (Fab)$'_2$ [sic]. Suitable production methods are described, for example, in Campbell, A.M., Monoclonal Antibody Technology, (1987) Elsevier Verlag, Amsterdam, New York, Oxford and in Breitling, F. and Dübel, S., Rekombinante Antikörper (1997), Spektrum Akademischer Verlag, Heidelberg.

Further use of the coding sequence:

The present cDNA additionally provides the basis for cloning the genomic sequence of the novel PARP gene. This also includes the relevant regulatory or promoter sequence, which is available, for example, by sequencing the region located 5′ upstream of the cDNA according to the invention. The cDNA sequence information is also the basis for producing antisense molecules or ribozymes with the aid of known methods (cf. Jones, J.T. and Sallenger, B.A. (1997) Nat. Biotechnol. 15, 902; Nellen, W. and Lichtenstein, C. (1993) TIBS, 18, 419). The genomic DNA can likewise be used to produce the gene constructs described above.

Another possibility of using the nucleotide sequence or parts thereof is to generate transgenic animals. Transgenic overexpression or genetic knock-out of the sequence information in suitable animal models may provide further valuable information about the (patho)physiology of the novel enzymes.

Therapeutic applications:

In situations where there is a prevailing deficiency of a protein according to the invention it is possible to employ several methods for replacement. On the one hand, the protein, natural or recombinant, can be administered directly or by gene therapy in the form of its coding nucleic acid (DNA or RNA). It is possible to use any suitable vectors for this, for example both viral and non-viral vehicles. Suitable methods are described, for example, by Strauss and Barranger in Concepts in Gene Therapy (1997), Walter de Gruyter, publisher. Another alternative is provided by stimulation of the endogenous gene by suitable agents.

It is also possible to block the turnover or the inactivation of PARPs according to the invention, for example by proteases. Finally, inhibitors or agonists of PARPs according to the invention can be employed.

In situations where a PARP is present in excess or is overactivated, various types of inhibitors can be employed. This inhibition can be achieved both by antisense molecules, ribozymes, oligonucleotides or antibodies, and by low molecular weight compounds.

Nontherapeutic applications:

The nucleic acids according to the invention, such as, for example, cDNA, the genomic DNA, the promoter, and the polypeptide, and partial fragments thereof, can also be used in recombinant or nonrecombinant form for developing various test systems.

For example, it is possible to establish a test system which is suitable for measuring the activity of the promoter or of the protein in the presence of a test substance. The methods of measurement in this case are preferably simple ones, e.g. colorimetric, luminometric, fluorimetric, immunological or radioactive, and allow preferably a large number of test substances to be measured rapidly. Tests of this type are suitable and advantageous for so-called high-throughput screening. These test systems allow test substances to be assessed for their binding to or their agonism, antagonism or inhibition of proteins according to the invention.

Determination of the amount, activity and distribution of the proteins according to the invention or their underlying mRNA in the human body can be used for the diagnosis, for the determination of the predisposition and for the monitoring of certain diseases. Likewise, the sequence of the cDNA and of [sic] the genomic sequence may provide information about genetic causes of and predispositions to certain diseases. It is possible to use for this purpose both DNA/RNA probes and antibodies of a wide variety of types. The nucleotide sequences according to the invention or parts thereof can further be used in the form of suitable probes for detecting point mutations, deletions or insertions.

The proteins according to the invention can further be used to identify and isolate their natural ligands or interacting partners. The proteins according to the invention can additionally be used to identify and isolate artificial or synthetic ligands. For this purpose, the recombinantly prepared or purified natural protein can be derivatized in such a way that it has modifications which permit linkage to support materials. Proteins bound in this way can be incubated with various analytes, such as, for example, protein extracts or peptide libraries or other sources of ligands. Specifically bound peptides, proteins or low molecular weight, non-proteinogenous substances can be isolated and characterized in this way. Non-proteinogenous substances mean, for example, low molecular weight chemical substances which may originate, for example, from

classical drug synthesis or from so-called substance libraries which have been synthesized combinatorially.

The protein extracts used are derived, for example, from homogenates of plants or parts of plants, microorganisms, human or animal tissues or organs.

Ligands or interacting partners can also be identified by methods like the yeast two-hybrid system (Fields, S. and Song, O. (1989) Nature, 340, 245). The expression banks which can be employed in this case may be derived, for example, from human tissues such as, for example, brain, heart, kidney etc.

The nucleic acid sequences according to the invention and the proteins encoded by them can be employed for developing reagents, agonists and antagonists or inhibitors for the diagnosis and therapy of chronic and acute diseases associated with the expression of one of the protein sequences according to the invention, such as, for example, with increased or decreased expression thereof. The reagents, agonists, antagonists or inhibitors developed can subsequently be used to produce pharmaceutical preparations for the treatment or diagnosis of disorders. Examples of possible diseases in this connection are those of the brain, of the peripheral nervous system, of the cardiovascular system or of the eye, of septic shock, of rheumatoid arthritis, diabetes, acute kidney failure, or of cancer.

The relevance of the proteins according to the invention for said indications was verified using specific inhibitors in relevant animal models (see examples). In a model of neurodegenrative [sic] disorders (NMDA excitotoxicity), the specific PARP2 inhibitor 2(4(2-(N,N-diethylamino)eth-1-yloxy)phenyl)-benzimidazole-4-carboxamide [sic] had surprisingly good activity (ED50 < 100 mg/kg).

The invention is now illustrated in detail with reference to the following examples.

Example 1: Isolation of the PARP2 and PARP3 cDNA

The present cDNA sequences were found for the first time on sequence analysis of cDNA clones of a cDNA library from human brain (Human Brain 5'Stretch Plus cDNA Library, # HL3002a, Clontech). The mouse PARP3 clones were isolated from a "ltriplex mouse brain cDNA library" (Clontech order No. ML5004t). The

sequences of these clones are described in SEQ ID NO:1, 3, 7 and 9.

Example 2: Expression of PARP2 and PARP3 in human tissues

The expression of human PARP2 and human PARP3 was investigated in eight different human tissues by Northern blot analysis. A Human Multiple Tissue Northern Blot (MTN™) supplied by Clontech (#7760-1 and #7780-1) was hybridized for this purpose with an RNA probe. The probe was produced by in vitro transcription of the corresponding cDNA of human PARP2 and human PARP3 in the presence of digoxigenin-labeled nucleotides in accordance with the manufacturer's method (BOEHRINGER MANNHEIM DIG Easy Hyb order No. 1603 558, DIG Easy Hyb method for RNA:RNA hybridization). The protocol was modified to carry out the prehybridization: 2x1h with addition of herring sperm DNA (10 mg/ml of hybridization solution). Hybridization then took place overnight with addition of herring sperm DNA (10 mg/ml of hybridization solution). The bands were detected using the CDP-Star protocol (BOEHRINGER MANNHEIM CDP-Star™ order No. 1685 627).

After stringent washing, the transcript of PARP2 was mainly detected in human brain, heart, skeletal muscle, kidney and liver. The transcript size of about 1.9 kb corresponds to the length of the cDNA determined (1.85kb) (cf. Figure 2(A)).

In other tissues or organs, human PARP2 expression is considerably weaker.

After stringent washing, the transcript of PARP3 was mainly expressed in heart, brain, kidney, skeletal muscle and liver. Expression in other tissues (placenta, lung, pancreas) is distinctly weaker (cf. Figure 2(B)). There are at least 2 transcripts for human PARP3, which can presumably be explained by different polyadenylation sites or alternative splicing. Their size (about 2.2 kb and 2.5 kb respectively) corresponds to the length of the cDNA determined (2.3kb). Washing was carried out with 0.2 SSC/0.2% SDS at room temperature for 2 x 15 minutes and then with 0.1 x SSC/0.1% SDS at 65°C for 2 x 15 minutes (prepared from 20X SSC: 3M NaCl, 0.3M sodium citrate, pH 7.0).

Example 3: Production of antibodies

Specific antibodies against the proteins according to the invention were produced. These were used inter alia for analyzing the tissue distribution at the protein level of PARP2 and PARP3 by

immunoblot (Western blot) analysis. Examples of the production of such antibodies are indicated below.

The following peptides were prepared by synthesis in the manner familiar to the skilled worker for the antibody production. In some cases, a cysteine residue was attached to the N or C terminals of the sequences in order to facilitate coupling to KLH (keyhole limpet hemocyanin).

PARP-2: NH_2-MAARRRRSTGGGRARALNES-CO_2H (amino acids 1-20)
NH_2-KTELQSPEHPLDQHYRNLHC-CO_2H (amino acids 335-353)
PARP-3: NH_2-CKGRQAGREEDPFRSTAEALK-CO_2H (amino acids 25-44)
NH_2-CKQQIARGFEALEALEEALK-CO_2H (amino acids 230-248)

The production of an anti-PARP3 antibody is described as a representative example.

For human PARP3, polyclonal antibodies were raised in rabbits using a synthetic peptide having the peptide sequence H_2N-KQQIARGFEALEALEEALK-CO_2H (amino acids 230-248 of the human PARP3 protein sequence). The corresponding mouse sequence differs in this region only by one amino acid (H_2N-KQQIARGFEALEALEEAMK-CO_2H). A cysteine was also attached to the N terminus in order to make it possible for the protein to couple to KLH.

Rabbits were immunized a total of five times, at intervals of 7-14 days, with the KLH-peptide conjugate. The antiserum obtained was affinity-purified using the antigen. The specific IgG fraction was isolated from the serum using the respective peptides which, for this purpose, were initially immobilized on an affinity column in the manner familiar to the skilled worker. The respective antiserum was loaded onto this affinity column, and nonspecifically sorbed proteins were eluted with buffer. The specifically bound IgG fraction was eluted with 0.2 M glycine/HCl buffer pH 2.2. The pH was immediately increased using a 1M TRIS/HCl buffer pH 7.5. The eluate containing the IgG fraction was mixed 1:1 (volume) with saturated ammonium sulfate solution and incubated at +4°C for 30 min to complete the precipitation. The resulting precipitate was centrifuged at 10,000 g and, after removal of the supernatant, dissolved in the minimum amount of PBS/TBS. The resulting solution was then dialyzed against PBS/TBS in the ratio 1:100 (volume). The antibodies were adjusted to a concentration of about 100 µg of IgG/ml. The PARP3 antibodies purified in this way had high specificity for PARP3. Whereas mouse PARP3 was recognized well, there was no observable cross-reaction with PARP1 or PARP2.

Example 4: Analysis of the tissue distribution by immunoblot (Western blot)

The tissue distribution at the protein level was also investigated for PARP2 and PARP3 by immunoblot (Western blot) analysis.

Preparation of the mouse tissues for protein gels:

Tissues or cells were homogenized using a Potter or Ultra-Turrax. For this, 0.5 g of tissue (or cells) was incubated in 5 ml of buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 6 mM $MgCl_2$) one tablet of protease inhibitor cocktail (Boehringer Mannheim, order No.: 1836153) and benzonase (purity grade I, MERCK) at 37°C for 30 min. Tissue samples from mice were produced for heart, lung, liver, spleen, kidney, intestine, muscle, brain and for human embryonic kidney cells (HEK293, human embryonal kidney).

Protein gels:

The NuPAGE system supplied by NOVEX was used according to the instructions for protein gels. Polyacrylamide gels (NuPAGE 4-12% BisTris, NOVEX NP 0321), running buffer (MES-Running Buffer, NOVEX NP 0002), antioxidant (NOVEX NP 0005), protein size standard (Multi Mark Multi Colored Standard, NOVEX LC 5725), sample buffer (NuPAGE LDS Sample Buffer (4X), NOVEX NP 0007) were used. The gels were run for 45 minutes at a voltage of 200 V.

Western blot:

Western blots were carried out using the NOVEX system in accordance with instructions. A nitrocellulose membrane (Nitrocellulose Pore size 45 µm, NOVEX LC 2001) was used. The transfer took 1 hour at a current of 200 mA. The transfer buffer consisted of 50 ml of transfer buffer concentrate (NOVEX NP 0006), 1 ml of antioxidant (NOVEX NP 0002), 100 ml of analytical grade methanol and 849 ml of double-distilled water.

Besides the blots produced in this way, also used were premade blots, for example from Chemicon (mouse brain blot, Chemicon, catalog No.: NS 106 with the tissues 1. frontal cortex, 2. posterior cortex, 3. cerebellum, 4. hippocampus, 5. olfactory bulb, 6. striatum, 7. thalamus, 8. mid brain, 9. entorhinal cortex, 10. pons, 11. medulla, 12. spinal cord).

Antibody reaction with PARP3:

The Western blots were blocked in TBST (TBS + 0.3 % Tween 20) with 5% dry milk powder for at least 2 hours (TBS: 100 mM Tris pH 7.5, 200 mM NaCl). The antibody reaction with the primary antibody (dilution 1:1000) took place in TBST with 5% dry milk powder (see above at room temperature for at least 2 hours or at 4°C overnight, with gentle agitation (vertical rotator). This was followed by washing three times in TBST for 5 minutes. Incubation with the secondary antibody (anti-rabbit IgG, peroxidase-coupled, SIGMA A-6154, dilution 1:2000) took place in TBST with 5% dry milk powder for 1 hour. This was followed by washing three times for 5 minutes each time as above. The subsequent detection was based on chemiluminescence using the SUPER BLAZE kit (Pierce, Signal BLAZE Chemiluminescent Substrate 34095) as stated by the manufacturer. The "Lumi-Film" (Chemiluminescent Detection Film, Boehringer order No: 1666916 were [sic] used. The films were developed for about 2 min (X-ray developer concentrate, ADEFO-Chemie GmbH), hydrated, fixed for about 4 min (Acidofix 85 g/l / AGFA), hydrated and then dried.

Example 5: Preparation of the enzymes

For comparison, human PARP1 was expressed recombinantly in the baculovirus system in the manner familiar to the skilled worker and partially purified as described (Shah et al., Analytical Biochemistry 1995, 227, 1-13). Bovine PARP1 in a purity of 30-50% (c= 0.22 mg/ml, spec. activity 170 nmol of ADP-ribose/min/mg of total protein at 25°C) was purchased from BIOMOL (order No. SE-165). Human and mouse PARP2 and PARP3 were expressed recombinantly in the baculovirus system (Bac-to-Bac system, BRL LifeScience). For this purpose, the appropriate cDNAs were cloned to the pFASTBAC-1 vector. Preparation of recombinant baculovirus DNA by recombination in E. coli was followed by transfection of insect cells (Sf9 or High-Five) with the appropriate recombinant baculovirus DNAs. Expression of the corresponding proteins was verified by Western blot analysis. Virus strains were amplified in the manner familiar to the skilled worker. Larger amounts of recombinant proteins were infected [sic] by infecting 500 ml of insect cell culture (2×10^6 cells/ml) with viruses in an MOI (multiplicity of infection; ratio of viruses to cells) of 5-10 and incubated for 3 to 4 days. The insect cells were then pelleted by centrifugation, and the proteins were purified from the pellet.

The purification took place by classical methods of protein purification familiar to the skilled worker, detecting the enzymes with appropriate specific antibodies. In some cases, the proteins were also affinity-purified on a 3-aminobenzamide affinity column as described (Burtscher et al., Anal Biochem 1986, 152:285-290). The purity was >90%.

Example 6: Assay systems for determining the from [sic] activity of PARP2 and PARP3 and the inhibitory action of effectors on PARP1, PARP2 and PARP3.

a) Production of antibodies against poly(ADP-ribose)

It is possible to use poly(ADP-ribose) as antigen for generating anti-poly(ADP-ribose) antibodies. The production of anti-poly(ADP-ribose) antibodies is described in the literature. (Kanai Y et al. (1974) Biochem Biophys Res Comm 59:1, 300-306; Kawamaitsu H et al. (1984) Biochemistry 23, 3771-3777; Kanai Y et al. (1978) Immunology 34, 501-508).

The following were used, inter alia: anti-poly(ADP-ribose) antibodies (polyclonal antiserum, rabbits), BIOMOL; order No. SA-276, anti-poly(ADP-ribose) antibodies (monoclonal, mouse; clone 10H; hybri-oma [sic] supernatant, affinity-purified).

The antisera or monoclonal antibodies obtained from hybridoma supernatant were purified by protein A affinity chromatography in the manner familiar to the skilled worker.

b) ELISA assay

Materials:

ELISA color reagent: TMB mix, SIGMA T-8540

A 96-well microtiter plate (FALCON Micro-Test IIIä Flexible Assay Plate, # 3912) was coated with histones (SIGMA, H-7755). Histones were for this purpose dissolved in carbonate buffer (0.05M Na_2HCO_3; pH 9.4) in a concentration of 50 µg/ml. The individual wells of the microtiter plate were each incubated with 150 µl of this histone solution at room temperature for at least 2 hours or at 4°C overnight. The wells are then blocked by adding 150 µl of a 1% strength BSA solution (SIGMA, A-7888) in carbonate buffer at room temperature for 2 hours. This is followed by three washing steps with washing buffer (0.05% Tween10 in 1x PBS; PBS (Phosphate buffered saline; Gibco, order No. 10010): 0.21g/l KH_2PO_4, 9g/l NaCl, 0.726g/l $Na_2HPO_4 \cdot 7H_2O$, pH 7.4). Washing steps were

all carried out in a microtiter plate washer ("Columbus" microtiter plate washer, SLT-Labinstruments, Austria).

Required for the enzyme reaction were an enzyme reaction solution and a substrate solution, in each case as a premix. The absolute amount of these solutions depended on the intended number of assay wells.

Composition of the enzyme reaction solution per well:
- 4 µl of PARP reaction buffer (1M Tris-HCl pH 8.0, 100mM $MgCl_2$, 10mM DTT)
- 20ng of PARP1 (human or bovine) or 8ng PARP2 (human or mouse)
- 4 µl of activated DNA (1 mg/ml; SIGMA, D-4522)
- H_2O ad 40 µl

Composition of the substrate solution per well:
- 5 µl of PARP reaction buffer (10x)
- 0.8 µl of NAD solution (10mM, SIGMA N-1511)
- 44 µl H_2O

Inhibitors were dissolved in 1x PARP reaction buffer. DMSO, that was occasionally used to dissolve inhibitors in higher concentrations, was no problem up to a final concentration of 2%. For the enzyme reaction, 40 µl of the enzyme reaction solution were introduced into each well and incubated with 10 µl of inhibitor solution for 10 minutes. The enzyme reaction was then started by adding 50 µl of substrate solution per well. The reaction was carried out at room temperature for 30 minutes and then stopped by washing three times with washing buffer.

The primary antibodies employed were specific anti-poly(ADP-ribose) antibodies in a dilution of 1:5000. Dilution took place in antibody buffer (1% BSA in PBS; 0.05% Tween20). The incubation time for the primary antibodies was one hour at room temperature. After subsequently washing three times with washing buffer, incubation was carried out with the secondary antibody (anti-mouse IgG, Fab fragments, peroxidase-coupled, Boehringer Mannheim, order No. 1500.686; anti-rabbit IgG, peroxidase-coupled, SIGMA, order No. A-6154) in a dilution of 1:10,000 in antibody buffer at room temperature for one hour. Washing three times with washing buffer was followed by the color reaction using 100 µl of color reagent (TMB mix, SIGMA) per well at room temperature for about 15 min. The color reaction was stopped by adding 100 µl of 2M H_2SO_4. This was followed by immediate measurement in an ELISA plate reader (EAR340AT "Easy Reader", SLT-Labinstruments, Aus-

tria) (450nm versus 620nm). The measurement principle is depicted diagrammatically in Figure 6.

Various concentrations were used to construct a dose-effect plot to determine the K_i value of an inhibitor. Values are obtained in triplicate for a particular inhibitor concentration. Arithmetic means are determined using Microsoft© Excel. The IC_{50} is determined using the Microcal© Origin Software (Vers. 5.0) ("Sigmoidal Fit"). Conversion of the IC_{50} value is calculated in this way into K_i values took place by using "calibration inhibitors". The "calibration inhibitors" were also measured in each analysis. The K_i values of the "calibration inhibitors" were determined in the same assay system by analysis of the Dixon diagram in the manner familiar to the skilled worker.

b) HTRF (homogenous time-resolved fluorescence) assay

In the HTFR [sic] PARP assay according to the invention, histones, as target proteins for modification by PARP, are labeled indirectly with an XL665 fluorophore. The antibody is directly labeled with a europium cryptate. If the XL665 fluorophore is in the direct vicinity in space, which is ensured by binding to the poly(ADP-ribose) on the histone, then energy transfer is possible. The emission at 665 nm is thus directly proportional to the amount of bound antibody, which in turn is equivalent to the amount of poly(ADP-ribose). The measured signal thus corresponds to the PARP activity. The measurement principle is depicted diagrammatically in Figure 7. The materials used are identical to those used in the ELISA assay (see above) unless expressly indicated.

Histones were dissolved in a concentration of 3 mg/ml in Hepes buffer (50mM, pH=7.5). Biotinylation took place with sulfo-NHS-LC-biotin (Pierce, #21335T). A molar ratio of 4 biotin per histone was used. The incubation time was 90 minutes (RT). The biotinylated histones were then purified on a G25 SF HR10/10 column (Pharmacia, 17-0591-01) in Hepes buffer (50mM, pH=7.0) in order to remove excess biotinylation reagent. The anti-poly(ADP-ribose) antibody was labeled with europium cryptate using bifunctional coupling reagents (Lopez, E. et al., Clin. Chem. 39(2), 196-201 (1993); US Patent 5,534,622). Purification took place on a G25SF HR10/30 column. A molar ratio of 3.1 cryptates per antibody was achieved. The yield was 25%. The conjugates were stored at -80°C in the presence of 0.1% BSA in phosphate buffer (0.1M, pH=7).

For the enzyme reaction, the following were pipetted into each

well:
- 10 µl of PARP solution in PARP HTRF reaction buffer (50mM Tris-HCl pH 8.0, 10mM MgCl2, 1mM DTT) with 20ng of PARP1 (human or bovine) or 8ng of PARP2 (human or mouse)
- 10 µl of activated DNA in PARP HTRF reaction buffer (50µg/ml)
- 10 µl of biotinylated histones in PARP HTRF reaction buffer (1.25µM)
- 10 µl of inhibitor in PARP HTRF reaction buffer

These reagents were incubated for 2 minutes before the reaction was started by adding
- 10 µl of NAD solution in PARP HTRF reaction buffer (41 µM/ml [sic]).

The reaction time was 30 minutes at room temperature.

The reaction was then stopped by adding
- 10 µl of PARP inhibitor (25 µM, K_i=10nM) in "Revelation" buffer (100mM Tris-HCl pH 7.2, 0.2M KF, 0.05% BSA).

The following were then added:
- 10 µl of EDTA solution (SIGMA, E-7889, 0.5M in H_2O)
- 100 µl of Sa-XL665 (Packard Instruments) in "Revelation" buffer (15-31.25nM)
- 50 µl of anti-PARP cryptate in "Revelation" buffer (1.6-3.3nM).

Measurement was then possible after 30 minutes (up to 4 hours). The measurement took place in a "discovery HTRF microplate analyzer" (Packard Instruments). The K_i values were calculated as described for the ELISA assay.

Example 7: Test systems for determining the therapeutic efficacy of PARP inhibitors

Novel PARP inhibitors can have their therapeutic efficacy checked in relevant pharmacological models. Examples of some suitable models are listed in Table 1.

Disorder	Model	Literature
Neurodegenerative disorders (stroke, Parkinson's, etc.)	NMDA excitotoxicity in mice or rats	See below for description

Stroke	Permanent MCAO ("middle cerebral artherial [sic] occlusion")	Tokime, T. et al., J. Cereb. Blood Flow Metab., 18(9): 991-7, 1998. Guegan, C., Brain Research. Molecular Brain Research, 55(1): 133-40, 1998.
	Transient, focal MCAO in rats or mice	Eliasson MJL et al., Nat Med 1997, 3:1089-1095. Endres, M et al., J Cereb Blood Flow Metab 1997, 17:1143-1151. Takahashi K et al., J Cereb Blood Flow Metab 1997, 17:1137-1142.
Parkinson's disease	MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) toxicity in mice/ rats	Cosi C, et al. , Brain Res. , 1998 809(1):58-67. Cosi C, et al. , Brain Res., 1996 729(2):264-9.
Myocardial infarct	Coronary vessel occlusion in rats, pigs or rabbits	Richard V, et al., Br. J. Pharmacol 1994, 113, 869-876. Thiemermann C, et al., Proc Natl Acad Sci U S A. 1997, 94(2):679-83. Zingarelli B, et al. , Cardiovasc Res. 1997, 36(2):205-15.
	Langendorf heart model in rats or rabbits	See below for des-cription
Septic shock	Endotoxin shock in rats	Szabo C, et al., J Clin Invest, 1997, 100(3):723-35.
	Zymosan- or carrageenan-induced multiple organ failure in rats or mice	Szabo C, et al. J Exp Med. 1997, 186(7):1041-9.Cuzzo-crea S, et al. Eur J Pharmacol. 1998, 342(1):67-76.
Rheumatoid arthritis	Adjuvant- or collagen-induced arthritis in rats or mice	Szabo C, et al., Proc Natl Acad Sci U S A. 1998, 95(7):3867-72.

Diabetes	Streptozotocin- and alloxan-induced or obesity-associated	Uchgata Y et al., Diabetes 1983, 32: 316-318.Masiello P et al., Diabetologia 1985, 28: 683-686.Shimabukuro M et al., J Clin Invest 1997, 100: 290-295.
Cancer	In vitro model; see below	Schlicker et al., 1999, 75(1), 91-100.

a) NMDA excitotoxicity model

Glutamate is the most important excitotory [sic] neurotransmitter in the brain. Under normal conditions, glutamate is secreted into the synaptic cleft and stimulates the post-synaptic glutamate receptors, specifically the glutamate receptors of the "NMDA" and "AMPA" types. This stimulation plays a significant part in numerous functions of the brain, including learning, memory and motor control.

Under the conditions of acute and chronic neurodegeneration (e.g. stroke), however, there is a great increase in the presynaptic glutamate secretion, resulting in excessive stimulation of the receptors. This leads to death of the cells stimulated in this way. These increased glutamate activities occur in a number of neurological disorders or psychological disturbances and lead to states of overexcitation or toxic effects in the central nervous system (CNS) but also in the peripheral nervous system. Thus, glutamate is involved in a large number of neurodegenerative disorders, in particular neurotoxic disturbances following hypoxia, anoxia, ischemia and after lesions like those occurring after stroke and trauma, and stroke, Alzheimer's disease, Huntington's disease, amyotrophic lateral sclerosis (ALS; "Lou Gehring's disease"), cranial trauma, spinal cord trauma, peripheral neuropathies, AIDS dementia and Parkinson's disease. Another disease in which glutamate receptors are important is epilepsy (cf. Brain Res Bull 1998; 46(4):281-309, Eur Neuropsychopharmacol 1998, 8(2):141-52.).

Glutamate effects are mediated through various receptors. One of these receptors is called the NMDA (N-methyl-D-aspartate) receptor after a specific agonist (Arzneim.Forschung 1990, 40, 511-514; TIPS, 1990, 11, 334-338; Drugs of the Future 1989, 14, 1059-1071). N-Methyl-D-aspartate is a strong agonist of a particular class of glutamate receptors ("NMDA" type). Stimulation of the NMDA receptor leads to influx of calcium into the cell and

the generation of free radicals. The free radicals lead to DNA damage and activation of PARP. PARP in turn causes cell death through depletion of high-energy phosphates (NAD and ATP) in the cell. This explains the toxicity of NMDA. Treatment of animals with NMDA can therefore be regarded as a model of the abovementioned disorders in which excitotoxicity is involved.

Because of the importance of glutamate receptors in neurodegeneration, many pharmacological approaches to date have been directed at specific blocking of precisely these receptors. However, because of their importance in normal stimulus conduction, these approaches have proved to be problematic (side effects). In addition, stimulation of the receptors is an event which takes place very rapidly so that administration of the receptors often comes too late ("time window" problem). Thus there is a great need for novel principles of action and inhibitors of NMDA-related neurotoxicity.

Protection against cerebral overexcitation by excitatory amino acids (NMDA antagonism in mice) can be regarded as adequate proof of the activity of a pharmacolical [sic] effector of PARP in disorders based on excitotoxicity. Intracerebral administration of excitatory amino acids (EAA) induces such massive overexcitation that it leads within a short time to convulsions and death of the animals (mice).

In the present case there was unilateral intracerebroventricular administration of 10 µl of a 0.035% strength aqueous NMDA solution 120 minutes after intraperitoneal (i.p.) administration of the test substance. These symptoms can be inhibited by systemic, e.g. intraperitoneal, administration of centrally acting drugs. Since excessive activation of EAA receptors in the central nervous system plays an important part in the pathogenesis of various neurological disorders, information can be gained from the detected EAA antagonism in vivo about possible therapeutic utilizability of the substances for such CNS disorders. An ED50 at which 50% of the animals are, due to preceding i.p. administration of the measured substance, free of syptoms [sic] with a fixed dose of either [sic] NMDA was determined as a measure of the activity of the substances.

The specific PARP2 inhibitor 2(4(2-(N,N-diethylamino)-eth-1-yloxy)phenyl)benzimidazole-4-carboxamide [sic] surprisingly shows an activity with an ED50 of about 30 mg/kg in this test.

b) Langendorf heart model (model for myocardial infarct)

Male Sprague-Dawley rats (bodyweight 300-400 g; origin Janvier, Le Genest-St-Isle, France) were used for the test. The rats were treated orally by gavage with the active substance or placebo (volume: 5 ml/kg). 50 minutes later, heparin is administered intraperitoneally (Liquemin N Roche, 125 IU/animal in 0.5 ml). The animals are anesthesized with Inactin® T133 (thiobetabarbital sodium 10%), fixed on the operating table, tracheotomized and ventilated with a "Harvard ventilatory pump" (40 beats/min, 4.5 ml/beat). Thoracotomy was followed by immediate catheterization of the aorta, removal of the heart and immediate retrograde perfusion. The hearts were perfused with a constant pressure of 75 mmHg, which is achieved using a "Gilson Miniplus 2 perfusion pump". Composition of the perfusate (mmol/l): NaCl 118, KCl 4.7, $CaCl_2$ x 2 H_2O 2.52, $MgSO_4$ x 7 H_2O 1.64, $NaHCO_3$ 24.88, KH_2PO_4 1.18, glucose 11. The temperature is kept at 37°C throughout the experiment. Functional parameters were continuously recorded using a "Gould 4-channel recorder". Measurements were made of the left-ventricular pressure (LVP; mmHg), LVEDP (mmHg), enzyme release (creatine kinase, mU/ml/g), coronary flow rate (ml/min), HR (pulse rate, min^{-1}). The left-ventricular pressure was measured using a liquid-filled latex balloon and a Statham23 Db pressure transducer. The volume of the balloon was initially adjusted to reach an LVEDP (left-venricular [sic] end-diastolic pressure) of about 12 mmHg. $Dpdt_{max}$ [sic] (maximum pumping force) is derived from the pressure signal using a differentiator module. The heart rate was calculated from the pressure signal. The flow rate was determined using a drop counter (BMT Messtechnik GmbH Berlin). After an equilibration time of 20 minutes, the hearts were subjected to a 30-minute global ischemia by stopping the perfusate supply while keeping the temperature at 37°C. During the following 60-minute reperfusion period, samples of the perfusate were taken after 3, 5, 10, 15, 30, 45 and 60 min for analysis of creatine kinase (CK) activity. Means and standard deviations for the measured parameters were analyzed statistically (Dunnett test). The significance limit was p=0.05.

The experiment on rabbit hearts was carried out similarly. Male white New Zealand rabbits (obtained from: Interfauna) were used. The hearts were prepared as described above for the rat model. The perfusion pressure was set at a maximum of 60 mmHg and the flow rate at about 25ml/min. The equilibration time was about 30 min. The substance was administered by infusion directly upstream of the heart. 15 min after starting the infusion, a 30-minute global ischemia was caused by stopping the flow while maintaining the temperature of the heart. A 30-minute reperfusion

followed. Perfusate was taken for investigation of CK activity before administration of the substance, after 15 min and at various times (5, 10, 15, 20, 30 min) during the reperfusion. The following parameters were measured: LVP (mmHg), LVEDP, LVdP/dt, PP (mmHg), HR (pulse rate; beats/min), CK activity (U/min/g heart weight).

c) Animal model for acute kidney failure

The protective effect of intravenous administration of PARP inhibitors (4 days) on the kidney function of rats with postischemic acute kidney failure was investigated.

Male Sprague-Dawley rats (about 330 g at the start of the experiments; breeder: Charles River) were used. 10-15 animals were employed per experimental group. Administration of active substance/placebo took place continuously with an osmotic micropump into the femoral vein. Orbital blood was taken (1.5 ml of whole blood) under inhalation anesthesia with enflurane (Ethrane Abbot, Wiesbaden).

After the initial measurements (blood sample) and determination of the amount of urine excreted in 24h, the rats were anesthetized ("Nembutal", pentobarbital sodium, Sanofi CEVA; 50mg/kg i.p., volume injected 1.0 ml/kg) and fastened on a heatable operating table (37°C). 125 IU/kg heparin (Liquemin N, Roche) were administered i.v. into the caudal vein. The abdominal cavity was opened and the right kidney was exposed. The branching-off renal artery was exposed and clamped off superiorly using bulldog clamps (Diefenbach 38mm). The left renal artery was likewise exposed and clamped off (superiorly, about half way to the kidney). During the operation, an osmotic micropump was implanted into the femoral vein. The intestine was reinserted and the fluid loss was compensated with luke-warm 0.9% NaCl. The animals were covered with a moist cloth and kept warm under red light. After 40 min, the appearance of the kidneys was recorded, and the clamps were removed, first the right then the left. The intestine was put back and 2 drops of antibiotic (Tardomyocel, Bayer) were add [sic]. The abdominal wall was closed with sterile cat gut (Ethicon No.4) and treated once more with 1 drop of antibiotic. The epidermis was sutured with sterile Ethibond Exel (Ethicon) No.3/0, and the suture was sprayed with Nebacetin N (Yamanouchi) wound spray. A tenth of a daily dose of drug/placebo is given as i.v. bolus.

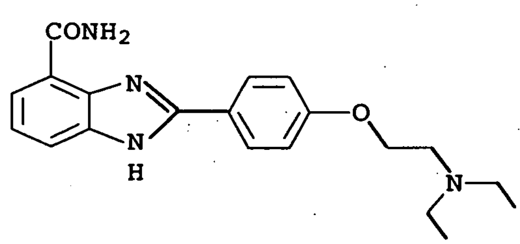
Samples and blood were taken for investigating biochemical parameters in the serum and urine: Na, K, creatinine, protein (only in urine), on days 1, 2 and 4 of the experiment. In addition, the feed and water consumption, bodyweight and urine volume were recorded. After 14 days, the animals were sacrificed and the kidneys were assessed.

The assessment excluded all animals which died of an infarct during the experiment or showed an infarct at necropsy on day 14. The creatinine clearance and the fractional sodium excretion were calculated as kidney function parameters, comparing treated animals with control and sham.

d) In vitro model for radiosensitization (tumor therapy)

MCF-7-cells (human breast carcinoma) were cultivated in Dulbecco's modified Eagle's medium with 10% heat-inactivated FCS and 2 mM L-glutamine. Cells were seeded out overnight in cell densities of 100, 1000 or 10,000 cells per well in a 6-well plate and then exposed to ionizing radiation with a dose in the range from 0 to 10 Gy (^{137}Cs, Shepard Mark, model [sic], model I-68A, dose rate 3.28 Gy/min). 10 days after the irradiation, the experiment was assessed, counting colonies with fifty cells as positive.

Example 8: Preparation of the PARP inhibitor 2(4(2-(N,N-diethylamino)eth-1-yloxy)phenyl)benzimidazole-4-carboxamide [sic]



a) 4(2-(N,N-Diethylaminoeth-1-yloxy)benzaldehyde [sic]

15 g (122 mmol) of 4-hydroxybenzaldehyde, 16.7 g (122 mmol) of N-(2-chloroethyl)-N,N-diethylamine and 33.9 g (246 mmol) of potassium carbonate were boiled under reflux together with a spatula tip of 18-crown-6 in 300 ml of ethyl methyl ketone for 6 hours. After filtration, the filtrate was concentrated in vacuo. The residue was partitioned between ether and 2M sodium hydroxide solution, and the ether phase was separated

off, dried and concentrated in vacuo. 24.8 g of the intermediate were obtained.

b) Ethyl 2(4(2-(N,N-diethylamino)eth-1-yloxy)phenyl)benzimidazole-4-carboxylate [sic]

2 g (11 mmol) of ethyl 2,3-diaminobenzoate and 1.4 ml of concentrated acetic acid were dissolved in 25 ml of methanol. Then 3.2 g (14.4 mmol) of the intermediate from stage a, dissolved in 50 ml of methanol, were added dropwise over the course of 30 minutes. Thereafter 2.9 g (14.4 mmol) of copper(II) acetate, dissolved in 37.5 ml of warm water, were rapidly added dropwise and then the mixture was boiled under reflux for 20 minutes. The reaction solution was cooled to 50°C, and 4.5 ml of 32% strength hydrochloric acid were added. Then a solution of 4.3 g of sodium sulfide hydrate in 25 ml of water was cautiously added dropwise, and the mixture was stirred for 15 minutes. The reaction solution was poured into ice-water, and the resulting precipitate was filtered off with suction. The filtrate was made alkaline with aqueous sodium bicarbonate solution and extracted several times with ethyl acetate. The ethyl acetate phase was separated off, dried and concentrated in vacuo. 4.4 g of the intermediate were obtained.

c) 2(4(2-(N,N-Diethylamino)eth-1-yloxy)phenyl)benzimidazole-4-carbohydrazide [sic]

2.7 g (54 mmol) of hydrazine hydrate were added to 4.1 g (10.7 mmol) of the intermediate from stage b in 30 ml of ethanol, and the mixture was boiled under reflux for 10 hours. The organic solvent was then removed in vacuo, and the residue was partitioned between water and ethyl acetate. The ethyl acetate phase was separated off, dried and concentrated in vacuo. The resulting residue was then treated with ether and again filtered with suction, resulting in 1.7 g of the intermediate.

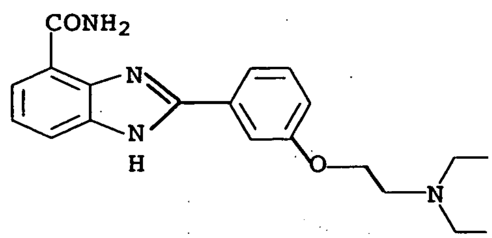
d) 2(4(2-(N,N-Diethylamino)eth-1-yloxy)phenyl)benzimidazole-4-carboxamide [sic]

About 1.6 g of Raney nickel were added to 1.6 g (4.5 mmol) of the intermediate from stage c in 45 ml of dimethylformamide/water (2/1), and the mixture was heated at 100°C for 6 hours. The reaction mixture was then filtered, and the filtrate was

diluted with a large amount of water, whereupon the product precipitated. 1.2 g of the product were obtained.

^{1}H-NMR (D_6-DMSO) δ = 0.95 (6H), 2.6 (4H), 2.8 (2H), 4.1 (2H), 7.1 (2H), 7.3 (1H), 7.7 (1H + NH), 7.85 (1H), 8.2 (2H) and 9.4 (NH) ppm.

Example 9: Preparation of the PARP inhibitor 2(3(2-(N,N-diethylamino)eth-1-yloxy)phenyl)benzimidazole-4-carboxamide [sic]



a) 3(2-(N,N-Diethylaminoeth-1-yloxy)benzaldehyde

6.1 g (50 mmol) of 3-hydroxybenzaldehyde were dissolved in 100 ml of ethanol, and 3.5 g (50 mmol) of sodium ethanolate were added. The mixture was stirred for 15 minutes. Then 7.5 g (55 mmol) of N-(2-chloroethyl)-N,N-diethylamine were added, and the mixture was boiled under reflux for 12 hours. The reaction mixture was then concentrated in vacuo. The residue was partitioned between ether and 1N sodium hydroxide solution, and the ether phase was separated off, dried and concentrated in vacuo. 7.6 g of the intermediate were obtained. [sic]

b) Ethyl 2(3(2-(N,N-diethylamino)eth-1-yloxy)phenyl)benzimidazole-4-carboxylate [sic]

1 g (5.5 mmol) of ethyl 2,3-diaminobenzoate and 0.68 ml of concentrated acetic acid were dissolved in 20 ml of methanol. Then 1.6 g (7.2 mmol) of the intermediate from stage a, dissolved in 30 ml of methanol, were added dropwise over the course of 30 minutes. Thereafter 1.1 g (5.5 mmol) of copper(II) acetate, dissolved in 19 ml of warm water, were rapidly added dropwise, and the mixture was then boiled under reflux for 20 minutes. The reaction solution was cooled to 50°C, and 2.25 ml of 32% strength hydrochloric acid were added. Then a solution of 2.13 g of sodium sulfide hydrate in

15 ml of water was cautiously added dropwise, and the mixture was stirred for 15 minutes. The reaction solution was poured into ice-water, and the resulting precipitate was filtered off with suction. The filtrate was made alkaline with aqueous sodium bicarbonate solution and extracted several times with ethyl acetate. The ethyl acetate phase was separated off, dried and concentrated in vacuo. 2.4 g of the intermediate were obtained.

c) 2(3(2-(N,N-Diethylamino)eth-1-yloxy)phenyl)benzimidazole-4-carbohydrazide [sic]

1.5 g (30 mmol) of hydrazine hydrate were added to 2.3 g (6.0 mmol) of the intermediate from stage b in 30 ml of butanol, and the mixture was heated at 120°C for 10 hours. The reaction mixture was then extracted with ethyl acetate. The ethyl acetate phase was separated off, dried and concentrated in vacuo. 1.7 g of the intermediate were obtained.

d) 2(3(2-N,N-Diethylamino)eth-1-yloxy)phenyl)benzimidazole-4-carboxamide [sic]

About 1.5 g of Raney nickel were added to 1 g (2.7 mmol) of the intermediate from stage c in 30 ml of dimethylformamide/water (2/1), and the mixture was heated at 100°C for 6 hours. The reaction mixture was then filtered, and the filtrate was diluted with a large amount of water, whereupon the product precipitated. 0.74 g of the product was obtained.

^{1}H-NMR (D_6-DMSO) d = 1.0 (6H), 2.6 (4H), 2.9 (2H), 4.15 (2H), 7.1 (1H), 7.4 (1H), 7.5 (1H), 7.7-7.9 (5H) and 9.3 (NH) ppm.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: BASF Aktiengesellschaft
(B) STREET: -
(C) CITY: Ludwigshafen
(E) COUNTRY: Germany
(F) POSTAL CODE: 67065

(ii) TITLE OF INVENTION: Novel poly(ADP-ribose) polymerase genes

(iii) NUMBER OF SEQUENCES: 10

(iv) COMPUTER-READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1843 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(vi) ORIGINAL SOURCE:
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..1715
(D) OTHER INFORMATION:/product= "Poly ADP Ribose Polymerase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CC ATG GCG GCG CGG CGG CGA CGG AGC ACC GGC GGC GGC AGG GCG AGA 47
Met Ala Ala Arg Arg Arg Arg Ser Thr Gly Gly Gly Arg Ala Arg
1 5 10 15

GCA TTA AAT GAA AGC AAA AGA GTT AAT AAT GGC AAC ACG GCT CCA GAA 95
Ala Leu Asn Glu Ser Lys Arg Val Asn Asn Gly Asn Thr Ala Pro Glu
20 25 30

GAC TCT TCC CCT GCC AAG AAA ACT CGT AGA TGC CAG AGA CAG GAG TCG 143
Asp Ser Ser Pro Ala Lys Lys Thr Arg Arg Cys Gln Arg Gln Glu Ser
35 40 45

AAA AAG ATG CCT GTG GCT GGA GGA AAA GCT AAT AAG GAC AGG ACA GAA 191
Lys Lys Met Pro Val Ala Gly Gly Lys Ala Asn Lys Asp Arg Thr Glu
50 55 60

GAC AAG CAA GAT GAA TCT GTG AAG GCC TTG CTG TTA AAG GGC AAA GCT 239
Asp Lys Gln Asp Glu Ser Val Lys Ala Leu Leu Leu Lys Gly Lys Ala
65 70 75

CCT GTG GAC CCA GAG TGT ACA GCC AAG GTG GGG AAG GCT CAT GTG TAT 287
Pro Val Asp Pro Glu Cys Thr Ala Lys Val Gly Lys Ala His Val Tyr
80 85 90 95

TGT GAA GGA AAT GAT GTC TAT GAT GTC ATG CTA AAT CAG ACC AAT CTC 335
Cys Glu Gly Asn Asp Val Tyr Asp Val Met Leu Asn Gln Thr Asn Leu
100 105 110

CAG TTC AAC AAC AAC AAG TAC TAT CTG ATT CAG CTA TTA GAA GAT GAT 383
Gln Phe Asn Asn Asn Lys Tyr Tyr Leu Ile Gln Leu Leu Glu Asp Asp
115 120 125

GCC CAG AGG AAC TTC AGT GTT TGG ATG AGA TGG GGC CGA GTT GGG AAA 431
Ala Gln Arg Asn Phe Ser Val Trp Met Arg Trp Gly Arg Val Gly Lys
130 135 140

ATG GGA CAG CAC AGC CTG GTG GCT TGT TCA GGC AAT CTC AAC AAG GCC 479
Met Gly Gln His Ser Leu Val Ala Cys Ser Gly Asn Leu Asn Lys Ala
145 150 155

AAG GAA ATC TTT CAG AAG AAA TTC CTT GAC AAA ACG AAA AAC AAT TGG 527
Lys Glu Ile Phe Gln Lys Lys Phe Leu Asp Lys Thr Lys Asn Asn Trp
160 165 170 175

GAA GAT CGA GAA AAG TTT GAG AAG GTG CCT GGA AAA TAT GAT ATG CTA 575

Glu Asp Arg Glu Lys Phe Glu Lys Val Pro Gly Lys Tyr Asp Met Leu
180 185 190

CAG ATG GAC TAT GCC ACC AAT ACT CAG GAT GAA GAG GAA ACA AAG AAA 623
Gln Met Asp Tyr Ala Thr Asn Thr Gln Asp Glu Glu Glu Thr Lys Lys
195 200 205

GAG GAA TCT CTT AAA TCT CCC TTG AAG CCA GAG TCA CAG CTA GAT CTT 671
Glu Glu Ser Leu Lys Ser Pro Leu Lys Pro Glu Ser Gln Leu Asp Leu
210 215 220

CGG GTA CAG GAG TTA ATA AAG TTG ATC TGT AAT GTT CAG GCC ATG GAA 719
Arg Val Gln Glu Leu Ile Lys Leu Ile Cys Asn Val Gln Ala Met Glu
225 230 235

GAA ATG ATG ATG GAA ATG AAG TAT AAT ACC AAG AAA GCC CCA CTT GGG 767
Glu Met Met Met Glu Met Lys Tyr Asn Thr Lys Lys Ala Pro Leu Gly
240 245 250 255

AAG CTG ACA GTG GCA CAA ATC AAG GCA GGT TAC CAG TCT CTT AAG AAG 815
Lys Leu Thr Val Ala Gln Ile Lys Ala Gly Tyr Gln Ser Leu Lys Lys
260 265 270

ATT GAG GAT TGT ATT CGG GCT GGC CAG CAT GGA CGA GCT CTC ATG GAA 863
Ile Glu Asp Cys Ile Arg Ala Gly Gln His Gly Arg Ala Leu Met Glu
275 280 285

GCA TGC AAT GAA TTC TAC ACC AGG ATT CCG CAT GAC TTT GGA CTC CGT 911
Ala Cys Asn Glu Phe Tyr Thr Arg Ile Pro His Asp Phe Gly Leu Arg
290 295 300

ACT CCT CCA CTA ATC CGG ACA CAG AAG GAA CTG TCA GAA AAA ATA CAA 959
Thr Pro Pro Leu Ile Arg Thr Gln Lys Glu Leu Ser Glu Lys Ile Gln
305 310 315

TTA CTA GAG GCT TTG GGA GAC ATT GAA ATT GCT ATT AAG CTG GTG AAA 1007
Leu Leu Glu Ala Leu Gly Asp Ile Glu Ile Ala Ile Lys Leu Val Lys
320 325 330 335

ACA GAG CTA CAA AGC CCA GAA CAC CCA TTG GAC CAA CAC TAT AGA AAC 1055
Thr Glu Leu Gln Ser Pro Glu His Pro Leu Asp Gln His Tyr Arg Asn
340 345 350

CTA CAT TGT GCC TTG CGC CCC CTT GAC CAT GAA AGT TAC GAG TTC AAA 1103
Leu His Cys Ala Leu Arg Pro Leu Asp His Glu Ser Tyr Glu Phe Lys
355 360 365

GTG ATT TCC CAG TAC CTA CAA TCT ACC CAT GCT CCC ACA CAC AGC GAC 1151
Val Ile Ser Gln Tyr Leu Gln Ser Thr His Ala Pro Thr His Ser Asp
370 375 380

TAT ACC ATG ACC TTG CTG GAT TTG TTT GAA GTG GAG AAG GAT GGT GAG 1199
Tyr Thr Met Thr Leu Leu Asp Leu Phe Glu Val Glu Lys Asp Gly Glu
385 390 395

AAA GAA GCC TTC AGA GAG GAC CTT CAT AAC AGG ATG CTT CTA TGG CAT 1247
Lys Glu Ala Phe Arg Glu Asp Leu His Asn Arg Met Leu Leu Trp His
400 405 410 415

GGT TCC AGG ATG AGT AAC TGG GTG GGA ATC TTG AGC CAT GGG CTT CGA 1295
Gly Ser Arg Met Ser Asn Trp Val Gly Ile Leu Ser His Gly Leu Arg
420 425 430

ATT GCC CCA CCT GAA GCT CCC ATC ACA GGT TAC ATG TTT GGG AAA GGA 1343
Ile Ala Pro Pro Glu Ala Pro Ile Thr Gly Tyr Met Phe Gly Lys Gly
435 440 445

ATC TAC TTT GCT GAC ATG TCT TCC AAG AGT GCC AAT TAC TGC TTT GCC 1391
Ile Tyr Phe Ala Asp Met Ser Ser Lys Ser Ala Asn Tyr Cys Phe Ala
450 455 460

TCT CGC CTA AAG AAT ACA GGA CTG CTG CTC TTA TCA GAG GTA GCT CTA 1439
Ser Arg Leu Lys Asn Thr Gly Leu Leu Leu Leu Ser Glu Val Ala Leu
465 470 475

GGT CAG TGT AAT GAA CTA CTA GAG GCC AAT CCT AAG GCC GAA GGA TTG 1487
Gly Gln Cys Asn Glu Leu Leu Glu Ala Asn Pro Lys Ala Glu Gly Leu
480 485 490 495

CTT CAA GGT AAA CAT AGC ACC AAG GGG CTG GGC AAG ATG GCT CCC AGT 1535
Leu Gln Gly Lys His Ser Thr Lys Gly Leu Gly Lys Met Ala Pro Ser
500 505 510

TCT GCC CAC TTC GTC ACC CTG AAT GGG AGT ACA GTG CCA TTA GGA CCA 1583
Ser Ala His Phe Val Thr Leu Asn Gly Ser Thr Val Pro Leu Gly Pro
515 520 525

GCA AGT GAC ACA GGA ATT CTG AAT CCA GAT GGT TAT ACC CTC AAC TAC 1631
Ala Ser Asp Thr Gly Ile Leu Asn Pro Asp Gly Tyr Thr Leu Asn Tyr
530 535 540

AAT GAA TAT ATT GTA TAT AAC CCC AAC CAG GTC CGT ATG CGG TAC CTT 1679
Asn Glu Tyr Ile Val Tyr Asn Pro Asn Gln Val Arg Met Arg Tyr Leu

545 550 555

TTA AAG GTT CAG TTT AAT TTC CTT CAG CTG TGG TGA ATGTTGATAT 1725
Leu Lys Val Gln Phe Asn Phe Leu Gln Leu Trp *
560 565 570

TAAATAAACC AGAGATCTGA TCTTCAAGCA AGAAAATAAG CAGTGTTGTA CTTGTGAATT 1785

TTGTGATATT TTATGTAATA AAAACTGTAC AGGTCTAAAA AAAAAAAAAA AAAAAAAA 1843

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 571 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Ala Arg Arg Arg Arg Ser Thr Gly Gly Gly Arg Ala Arg Ala
1 5 10 15

Leu Asn Glu Ser Lys Arg Val Asn Asn Gly Asn Thr Ala Pro Glu Asp
20 25 30

Ser Ser Pro Ala Lys Lys Thr Arg Arg Cys Gln Arg Gln Glu Ser Lys
35 40 45

Lys Met Pro Val Ala Gly Gly Lys Ala Asn Lys Asp Arg Thr Glu Asp
50 55 60

Lys Gln Asp Glu Ser Val Lys Ala Leu Leu Leu Lys Gly Lys Ala Pro
65 70 75 80

Val Asp Pro Glu Cys Thr Ala Lys Val Gly Lys Ala His Val Tyr Cys
85 90 95

Glu Gly Asn Asp Val Tyr Asp Val Met Leu Asn Gln Thr Asn Leu Gln
100 105 110

Phe Asn Asn Asn Lys Tyr Tyr Leu Ile Gln Leu Leu Glu Asp Asp Ala
115 120 125

Gln Arg Asn Phe Ser Val Trp Met Arg Trp Gly Arg Val Gly Lys Met
130 135 140

```
Gly Gln His Ser Leu Val Ala Cys Ser Gly Asn Leu Asn Lys Ala Lys
145                 150                 155                 160

Glu Ile Phe Gln Lys Lys Phe Leu Asp Lys Thr Lys Asn Asn Trp Glu
                165                 170                 175

Asp Arg Glu Lys Phe Glu Lys Val Pro Gly Lys Tyr Asp Met Leu Gln
            180                 185                 190

Met Asp Tyr Ala Thr Asn Thr Gln Asp Glu Glu Glu Thr Lys Lys Glu
        195                 200                 205

Glu Ser Leu Lys Ser Pro Leu Lys Pro Glu Ser Gln Leu Asp Leu Arg
    210                 215                 220

Val Gln Glu Leu Ile Lys Leu Ile Cys Asn Val Gln Ala Met Glu Glu
225                 230                 235                 240

Met Met Met Glu Met Lys Tyr Asn Thr Lys Lys Ala Pro Leu Gly Lys
                245                 250                 255

Leu Thr Val Ala Gln Ile Lys Ala Gly Tyr Gln Ser Leu Lys Lys Ile
            260                 265                 270

Glu Asp Cys Ile Arg Ala Gly Gln His Gly Arg Ala Leu Met Glu Ala
        275                 280                 285

Cys Asn Glu Phe Tyr Thr Arg Ile Pro His Asp Phe Gly Leu Arg Thr
    290                 295                 300

Pro Pro Leu Ile Arg Thr Gln Lys Glu Leu Ser Glu Lys Ile Gln Leu
305                 310                 315                 320

Leu Glu Ala Leu Gly Asp Ile Glu Ile Ala Ile Lys Leu Val Lys Thr
                325                 330                 335

Glu Leu Gln Ser Pro Glu His Pro Leu Asp Gln His Tyr Arg Asn Leu
            340                 345                 350

His Cys Ala Leu Arg Pro Leu Asp His Glu Ser Tyr Glu Phe Lys Val
        355                 360                 365

Ile Ser Gln Tyr Leu Gln Ser Thr His Ala Pro Thr His Ser Asp Tyr
    370                 375                 380

Thr Met Thr Leu Leu Asp Leu Phe Glu Val Glu Lys Asp Gly Glu Lys
```

385 390 395 400

Glu Ala Phe Arg Glu Asp Leu His Asn Arg Met Leu Leu Trp His Gly
405 410 415

Ser Arg Met Ser Asn Trp Val Gly Ile Leu Ser His Gly Leu Arg Ile
420 425 430

Ala Pro Pro Glu Ala Pro Ile Thr Gly Tyr Met Phe Gly Lys Gly Ile
435 440 445

Tyr Phe Ala Asp Met Ser Ser Lys Ser Ala Asn Tyr Cys Phe Ala Ser
450 455 460

Arg Leu Lys Asn Thr Gly Leu Leu Leu Leu Ser Glu Val Ala Leu Gly
465 470 475 480

Gln Cys Asn Glu Leu Leu Glu Ala Asn Pro Lys Ala Glu Gly Leu Leu
485 490 495

Gln Gly Lys His Ser Thr Lys Gly Leu Gly Lys Met Ala Pro Ser Ser
500 505 510

Ala His Phe Val Thr Leu Asn Gly Ser Thr Val Pro Leu Gly Pro Ala
515 520 525

Ser Asp Thr Gly Ile Leu Asn Pro Asp Gly Tyr Thr Leu Asn Tyr Asn
530 535 540

Glu Tyr Ile Val Tyr Asn Pro Asn Gln Val Arg Met Arg Tyr Leu Leu
545 550 555 560

Lys Val Gln Phe Asn Phe Leu Gln Leu Trp *
565 570

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2265 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(vi) ORIGINAL SOURCE:
(F) TISSUE TYPE: uterus

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 242..1843
(D) OTHER INFORMATION:/product= "Poly ADP Ribose Polymerase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
TGGGACTGGT CGCCTGACTC GGCCTGCCCC AGCCTCTGCT TCACCCCACT GGTGGCCAAA      60

TAGCCGATGT CTAATCCCCC ACACAAGCTC ATCCCCGGCC TCTGGGATTG TTGGGAATTC     120

TCTCCCTAAT TCACGCCTGA GGCTCATGGA GAGTTGCTAG ACCTGGGACT GCCCTGGGAG     180

GCGCACACAA CCAGGCCGGG TGGCAGCCAG GACCTCTCCC ATGTCCCTGC TTTTCTTGGC     240

C ATG GCT CCA AAG CCG AAG CCC TGG GTA CAG ACT GAG GGC CCT GAG         286
  Met Ala Pro Lys Pro Lys Pro Trp Val Gln Thr Glu Gly Pro Glu
              575                 580                 585

AAG AAG AAG GGC CGG CAG GCA GGA AGG GAG GAG GAC CCC TTC CGC TCC       334
Lys Lys Lys Gly Arg Gln Ala Gly Arg Glu Glu Asp Pro Phe Arg Ser
            590                 595                 600

ACC GCT GAG GCC CTC AAG GCC ATA CCC GCA GAG AAG CGC ATA ATC CGC       382
Thr Ala Glu Ala Leu Lys Ala Ile Pro Ala Glu Lys Arg Ile Ile Arg
        605                 610                 615

GTG GAT CCA ACA TGT CCA CTC AGC AGC AAC CCC GGG ACC CAG GTG TAT       430
Val Asp Pro Thr Cys Pro Leu Ser Ser Asn Pro Gly Thr Gln Val Tyr
    620                 625                 630

GAG GAC TAC AAC TGC ACC CTG AAC CAG ACC AAC ATC GAG AAC AAC AAC       478
Glu Asp Tyr Asn Cys Thr Leu Asn Gln Thr Asn Ile Glu Asn Asn Asn
635                 640                 645                 650

AAC AAG TTC TAC ATC ATC CAG CTG CTC CAA GAC AGC AAC CGC TTC TTC       526
Asn Lys Phe Tyr Ile Ile Gln Leu Leu Gln Asp Ser Asn Arg Phe Phe
                655                 660                 665

ACC TGC TGG AAC CGC TGG GGC CGT GTG GGA GAG GTC GGC CAG TCA AAG       574
Thr Cys Trp Asn Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys
```

670 675 680

ATC AAC CAC TTC ACA AGG CTA GAA GAT GCA AAG AAG GAC TTT GAG AAG 622
Ile Asn His Phe Thr Arg Leu Glu Asp Ala Lys Lys Asp Phe Glu Lys
685 690 695

AAA TTT CGG GAA AAG ACC AAG AAC AAC TGG GCA GAG CGG GAC CAC TTT 670
Lys Phe Arg Glu Lys Thr Lys Asn Asn Trp Ala Glu Arg Asp His Phe
700 705 710

GTG TCT CAC CCG GGC AAG TAC ACA CTT ATC GAA GTA CAG GCA GAG GAT 718
Val Ser His Pro Gly Lys Tyr Thr Leu Ile Glu Val Gln Ala Glu Asp
715 720 725 730

GAG GCC CAG GAA GCT GTG GTG AAG GTG GAC AGA GGC CCA GTG AGG ACT 766
Glu Ala Gln Glu Ala Val Val Lys Val Asp Arg Gly Pro Val Arg Thr
735 740 745

GTG ACT AAG CGG GTG CAG CCC TGC TCC CTG GAC CCA GCC ACG CAG AAG 814
Val Thr Lys Arg Val Gln Pro Cys Ser Leu Asp Pro Ala Thr Gln Lys
750 755 760

CTC ATC ACT AAC ATC TTC AGC AAG GAG ATG TTC AAG AAC ACC ATG GCC 862
Leu Ile Thr Asn Ile Phe Ser Lys Glu Met Phe Lys Asn Thr Met Ala
765 770 775

CTC ATG GAC CTG GAT GTG AAG AAG ATG CCC CTG GGA AAG CTG AGC AAG 910
Leu Met Asp Leu Asp Val Lys Lys Met Pro Leu Gly Lys Leu Ser Lys
780 785 790

CAA CAG ATT GCA CGG GGT TTC GAG GCC TTG GAG GCG CTG GAG GAG GCC 958
Gln Gln Ile Ala Arg Gly Phe Glu Ala Leu Glu Ala Leu Glu Glu Ala
795 800 805 810

CTG AAA GGC CCC ACG GAT GGT GGC CAA AGC CTG GAG GAG CTG TCC TCA 1006
Leu Lys Gly Pro Thr Asp Gly Gly Gln Ser Leu Glu Glu Leu Ser Ser
815 820 825

CAC TTT TAC ACC GTC ATC CCG CAC AAC TTC GGC CAC AGC CAG CCC CCG 1054
His Phe Tyr Thr Val Ile Pro His Asn Phe Gly His Ser Gln Pro Pro
830 835 840

CCC ATC AAT TCC CCT GAG CTT CTG CAG GCC AAG AAG GAC ATG CTG CTG 1102
Pro Ile Asn Ser Pro Glu Leu Leu Gln Ala Lys Lys Asp Met Leu Leu
845 850 855

GTG CTG GCG GAC ATC GAG CTG GCC CAG GCC CTG CAG GCA GTC TCT GAG 1150

Val Leu Ala Asp Ile Glu Leu Ala Gln Ala Leu Gln Ala Val Ser Glu
    860   865   870

CAG GAG AAG ACG GTG GAG GAG GTG CCA CAC CCC CTG GAC CGA GAC TAC 1198
Gln Glu Lys Thr Val Glu Glu Val Pro His Pro Leu Asp Arg Asp Tyr
875   880   885   890

CAG CTT CTC AAG TGC CAG CTG CAG CTG CTA GAC TCT GGA GCA CCT GAG 1246
Gln Leu Leu Lys Cys Gln Leu Gln Leu Leu Asp Ser Gly Ala Pro Glu
                895   900   905

TAC AAG GTG ATA CAG ACC TAC TTA GAA CAG ACT GGC AGC AAC CAC AGG 1294
Tyr Lys Val Ile Gln Thr Tyr Leu Glu Gln Thr Gly Ser Asn His Arg
            910   915   920

TGC CCT ACA CTT CAA CAC ATC TGG AAA GTA AAC CAA GAA GGG GAG GAA 1342
Cys Pro Thr Leu Gln His Ile Trp Lys Val Asn Gln Glu Gly Glu Glu
        925   930   935

GAC AGA TTC CAG GCC CAC TCC AAA CTG GGT AAT CGG AAG CTG CTG TGG 1390
Asp Arg Phe Gln Ala His Ser Lys Leu Gly Asn Arg Lys Leu Leu Trp
    940   945   950

CAT GGC ACC AAC ATG GCC GTG GTG GCC GCC ATC CTC ACT AGT GGG CTC 1438
His Gly Thr Asn Met Ala Val Val Ala Ala Ile Leu Thr Ser Gly Leu
955   960   965   970

CGC ATC ATG CCA CAT TCT GGT GGG CGT GTT GGC AAG GGC ATC TAC TTT 1486
Arg Ile Met Pro His Ser Gly Gly Arg Val Gly Lys Gly Ile Tyr Phe
                975   980   985

GCC TCA GAG AAC AGC AAG TCA GCT GGA TAT GTT ATT GGC ATG AAG TGT 1534
Ala Ser Glu Asn Ser Lys Ser Ala Gly Tyr Val Ile Gly Met Lys Cys
            990   995   1000

GGG GCC CAC CAT GTC GGC TAC ATG TTC CTG GGT GAG GTG GCC CTG GGC 1582
Gly Ala His His Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu Gly
        1005   1010   1015

AGA GAG CAC CAT ATC AAC ACG GAC AAC CCC AGC TTG AAG AGC CCA CCT 1630
Arg Glu His His Ile Asn Thr Asp Asn Pro Ser Leu Lys Ser Pro Pro
    1020   1025   1030

CCT GGC TTC GAC AGT GTC ATT GCC CGA GGC CAC ACC GAG CCT GAT CCG 1678
Pro Gly Phe Asp Ser Val Ile Ala Arg Gly His Thr Glu Pro Asp Pro
1035   1040   1045   1050

ACC CAG GAC ACT GAG TTG GAG CTG GAT GGC CAG CAA GTG GTG GTG CCC 1726
Thr Gln Asp Thr Glu Leu Glu Leu Asp Gly Gln Gln Val Val Val Pro
1055 1060 1065

CAG GGC CAG CCT GTG CCC TGC CCA GAG TTC AGC AGC TCC ACA TTC TCC 1774
Gln Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe Ser
1070 1075 1080

CAG AGC GAG TAC CTC ATC TAC CAG GAG AGC CAG TGT CGC CTG CGC TAC 1822
Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg Tyr
1085 1090 1095

CTG CTG GAG GTC CAC CTC TGA GTGCCCGCCC TGTCCCCCGG GGTCCTGCAA 1873
Leu Leu Glu Val His Leu *
1100 1105

GGCTGGACTG TGATCTTCAA TCATCCTGCC CATCTCTGGT ACCCCTATAT CACTCCTTTT 1933

TTTCAAGAAT ACAATACGTT GTTGTTAACT ATAGTCACCA TGCTGTACAA GATCCCTGAA 1993

CTTATGCCTC CTAACTGAAA TTTTGTATTC TTTGACACAT CTGCCCAGTC CCTCTCCTCC 2053

CAGCCCATGG TAACCAGCAT TTGACTCTTT ACTTGTATAA GGGCAGCTTT TATAGGTTCC 2113

ACATGTAAGT GAGATCATGC AGTGTTTGTC TTTCTGTGCC TGGCTTATTT CACTCAGCAT 2173

AATGTGCACC GGGTTCACCC ATGTTTTCAT AAATGACAAG ATTTCCTCCT TTAAAAAAAA 2233

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA 2265

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 534 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Ala Pro Lys Pro Lys Pro Trp Val Gln Thr Glu Gly Pro Glu Lys
1 5 10 15

Lys Lys Gly Arg Gln Ala Gly Arg Glu Glu Asp Pro Phe Arg Ser Thr
20 25 30

Ala Glu Ala Leu Lys Ala Ile Pro Ala Glu Lys Arg Ile Ile Arg Val

35 40 45

Asp Pro Thr Cys Pro Leu Ser Ser Asn Pro Gly Thr Gln Val Tyr Glu
50 55 60

Asp Tyr Asn Cys Thr Leu Asn Gln Thr Asn Ile Glu Asn Asn Asn Asn
65 70 75 80

Lys Phe Tyr Ile Ile Gln Leu Leu Gln Asp Ser Asn Arg Phe Phe Thr
85 90 95

Cys Trp Asn Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys Ile
100 105 110

Asn His Phe Thr Arg Leu Glu Asp Ala Lys Lys Asp Phe Glu Lys Lys
115 120 125

Phe Arg Glu Lys Thr Lys Asn Asn Trp Ala Glu Arg Asp His Phe Val
130 135 140

Ser His Pro Gly Lys Tyr Thr Leu Ile Glu Val Gln Ala Glu Asp Glu
145 150 155 160

Ala Gln Glu Ala Val Val Lys Val Asp Arg Gly Pro Val Arg Thr Val
165 170 175

Thr Lys Arg Val Gln Pro Cys Ser Leu Asp Pro Ala Thr Gln Lys Leu
180 185 190

Ile Thr Asn Ile Phe Ser Lys Glu Met Phe Lys Asn Thr Met Ala Leu
195 200 205

Met Asp Leu Asp Val Lys Lys Met Pro Leu Gly Lys Leu Ser Lys Gln
210 215 220

Gln Ile Ala Arg Gly Phe Glu Ala Leu Glu Ala Leu Glu Glu Ala Leu
225 230 235 240

Lys Gly Pro Thr Asp Gly Gly Gln Ser Leu Glu Glu Leu Ser Ser His
245 250 255

Phe Tyr Thr Val Ile Pro His Asn Phe Gly His Ser Gln Pro Pro Pro
260 265 270

Ile Asn Ser Pro Glu Leu Leu Gln Ala Lys Lys Asp Met Leu Leu Val
275 280 285

Leu Ala Asp Ile Glu Leu Ala Gln Ala Leu Gln Ala Val Ser Glu Gln
290 295 300

Glu Lys Thr Val Glu Glu Val Pro His Pro Leu Asp Arg Asp Tyr Gln
305 310 315 320

Leu Leu Lys Cys Gln Leu Gln Leu Leu Asp Ser Gly Ala Pro Glu Tyr
325 330 335

Lys Val Ile Gln Thr Tyr Leu Glu Gln Thr Gly Ser Asn His Arg Cys
340 345 350

Pro Thr Leu Gln His Ile Trp Lys Val Asn Gln Glu Gly Glu Glu Asp
355 360 365

Arg Phe Gln Ala His Ser Lys Leu Gly Asn Arg Lys Leu Leu Trp His
370 375 380

Gly Thr Asn Met Ala Val Val Ala Ala Ile Leu Thr Ser Gly Leu Arg
385 390 395 400

Ile Met Pro His Ser Gly Gly Arg Val Gly Lys Gly Ile Tyr Phe Ala
405 410 415

Ser Glu Asn Ser Lys Ser Ala Gly Tyr Val Ile Gly Met Lys Cys Gly
420 425 430

Ala His His Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu Gly Arg
435 440 445

Glu His His Ile Asn Thr Asp Asn Pro Ser Leu Lys Ser Pro Pro Pro
450 455 460

Gly Phe Asp Ser Val Ile Ala Arg Gly His Thr Glu Pro Asp Pro Thr
465 470 475 480

Gln Asp Thr Glu Leu Glu Leu Asp Gly Gln Gln Val Val Val Pro Gln
485 490 495

Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe Ser Gln
500 505 510

Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg Tyr Leu
515 520 525

Leu Glu Val His Leu *
530

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2265 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(vi) ORIGINAL SOURCE:
(F) TISSUE TYPE: uterus

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 221..1843
(D) OTHER INFORMATION:/product= "Poly ADP Ribose Polymerase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
TGGGACTGGT CGCCTGACTC GGCCTGCCCC AGCCTCTGCT TCACCCCACT GGTGGCCAAA      60

TAGCCGATGT CTAATCCCCC ACACAAGCTC ATCCCCGGCC TCTGGGATTG TTGGGAATTC     120

TCTCCCTAAT TCACGCCTGA GGCTCATGGA GAGTTGCTAG ACCTGGGACT GCCCTGGGAG     180

GCGCACACAA CCAGGCCGGG TGGCAGCCAG GACCTCTCCC ATG TCC CTG CTT TTC       235
                                            Met Ser Leu Leu Phe
                                            535

TTG GCC ATG GCT CCA AAG CCG AAG CCC TGG GTA CAG ACT GAG GGC CCT       283
Leu Ala Met Ala Pro Lys Pro Lys Pro Trp Val Gln Thr Glu Gly Pro
540                 545                 550                 555

GAG AAG AAG AAG GGC CGG CAG GCA GGA AGG GAG GAG GAC CCC TTC CGC       331
Glu Lys Lys Lys Gly Arg Gln Ala Gly Arg Glu Glu Asp Pro Phe Arg
                560                 565                 570

TCC ACC GCT GAG GCC CTC AAG GCC ATA CCC GCA GAG AAG CGC ATA ATC       379
Ser Thr Ala Glu Ala Leu Lys Ala Ile Pro Ala Glu Lys Arg Ile Ile
            575                 580                 585
```

CGC GTG GAT CCA ACA TGT CCA CTC AGC AGC AAC CCC GGG ACC CAG GTG 427
Arg Val Asp Pro Thr Cys Pro Leu Ser Ser Asn Pro Gly Thr Gln Val
590 595 600

TAT GAG GAC TAC AAC TGC ACC CTG AAC CAG ACC AAC ATC GAG AAC AAC 475
Tyr Glu Asp Tyr Asn Cys Thr Leu Asn Gln Thr Asn Ile Glu Asn Asn
605 610 615

AAC AAC AAG TTC TAC ATC ATC CAG CTG CTC CAA GAC AGC AAC CGC TTC 523
Asn Asn Lys Phe Tyr Ile Ile Gln Leu Leu Gln Asp Ser Asn Arg Phe
620 625 630 635

TTC ACC TGC TGG AAC CGC TGG GGC CGT GTG GGA GAG GTC GGC CAG TCA 571
Phe Thr Cys Trp Asn Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser
640 645 650

AAG ATC AAC CAC TTC ACA AGG CTA GAA GAT GCA AAG AAG GAC TTT GAG 619
Lys Ile Asn His Phe Thr Arg Leu Glu Asp Ala Lys Lys Asp Phe Glu
655 660 665

AAG AAA TTT CGG GAA AAG ACC AAG AAC AAC TGG GCA GAG CGG GAC CAC 667
Lys Lys Phe Arg Glu Lys Thr Lys Asn Asn Trp Ala Glu Arg Asp His
670 675 680

TTT GTG TCT CAC CCG GGC AAG TAC ACA CTT ATC GAA GTA CAG GCA GAG 715
Phe Val Ser His Pro Gly Lys Tyr Thr Leu Ile Glu Val Gln Ala Glu
685 690 695

GAT GAG GCC CAG GAA GCT GTG GTG AAG GTG GAC AGA GGC CCA GTG AGG 763
Asp Glu Ala Gln Glu Ala Val Val Lys Val Asp Arg Gly Pro Val Arg
700 705 710 715

ACT GTG ACT AAG CGG GTG CAG CCC TGC TCC CTG GAC CCA GCC ACG CAG 811
Thr Val Thr Lys Arg Val Gln Pro Cys Ser Leu Asp Pro Ala Thr Gln
720 725 730

AAG CTC ATC ACT AAC ATC TTC AGC AAG GAG ATG TTC AAG AAC ACC ATG 859
Lys Leu Ile Thr Asn Ile Phe Ser Lys Glu Met Phe Lys Asn Thr Met
735 740 745

GCC CTC ATG GAC CTG GAT GTG AAG AAG ATG CCC CTG GGA AAG CTG AGC 907
Ala Leu Met Asp Leu Asp Val Lys Lys Met Pro Leu Gly Lys Leu Ser
750 755 760

AAG CAA CAG ATT GCA CGG GGT TTC GAG GCC TTG GAG GCG CTG GAG GAG 955
Lys Gln Gln Ile Ala Arg Gly Phe Glu Ala Leu Glu Ala Leu Glu Glu

```
    765                 770                 775

GCC CTG AAA GGC CCC ACG GAT GGT GGC CAA AGC CTG GAG GAG CTG TCC      1003
Ala Leu Lys Gly Pro Thr Asp Gly Gly Gln Ser Leu Glu Glu Leu Ser
780                 785                 790                 795

TCA CAC TTT TAC ACC GTC ATC CCG CAC AAC TTC GGC CAC AGC CAG CCC      1051
Ser His Phe Tyr Thr Val Ile Pro His Asn Phe Gly His Ser Gln Pro
                800                 805                 810

CCG CCC ATC AAT TCC CCT GAG CTT CTG CAG GCC AAG AAG GAC ATG CTG      1099
Pro Pro Ile Asn Ser Pro Glu Leu Leu Gln Ala Lys Lys Asp Met Leu
            815                 820                 825

CTG GTG CTG GCG GAC ATC GAG CTG GCC CAG GCC CTG CAG GCA GTC TCT      1147
Leu Val Leu Ala Asp Ile Glu Leu Ala Gln Ala Leu Gln Ala Val Ser
        830                 835                 840

GAG CAG GAG AAG ACG GTG GAG GAG GTG CCA CAC CCC CTG GAC CGA GAC      1195
Glu Gln Glu Lys Thr Val Glu Glu Val Pro His Pro Leu Asp Arg Asp
    845                 850                 855

TAC CAG CTT CTC AAG TGC CAG CTG CAG CTG CTA GAC TCT GGA GCA CCT      1243
Tyr Gln Leu Leu Lys Cys Gln Leu Gln Leu Leu Asp Ser Gly Ala Pro
860                 865                 870                 875

GAG TAC AAG GTG ATA CAG ACC TAC TTA GAA CAG ACT GGC AGC AAC CAC      1291
Glu Tyr Lys Val Ile Gln Thr Tyr Leu Glu Gln Thr Gly Ser Asn His
                880                 885                 890

AGG TGC CCT ACA CTT CAA CAC ATC TGG AAA GTA AAC CAA GAA GGG GAG      1339
Arg Cys Pro Thr Leu Gln His Ile Trp Lys Val Asn Gln Glu Gly Glu
            895                 900                 905

GAA GAC AGA TTC CAG GCC CAC TCC AAA CTG GGT AAT CGG AAG CTG CTG      1387
Glu Asp Arg Phe Gln Ala His Ser Lys Leu Gly Asn Arg Lys Leu Leu
        910                 915                 920

TGG CAT GGC ACC AAC ATG GCC GTG GTG GCC GCC ATC CTC ACT AGT GGG      1435
Trp His Gly Thr Asn Met Ala Val Val Ala Ala Ile Leu Thr Ser Gly
    925                 930                 935

CTC CGC ATC ATG CCA CAT TCT GGT GGG CGT GTT GGC AAG GGC ATC TAC      1483
Leu Arg Ile Met Pro His Ser Gly Gly Arg Val Gly Lys Gly Ile Tyr
940                 945                 950                 955

TTT GCC TCA GAG AAC AGC AAG TCA GCT GGA TAT GTT ATT GGC ATG AAG      1531
```

Phe Ala Ser Glu Asn Ser Lys Ser Ala Gly Tyr Val Ile Gly Met Lys
960 965 970

TGT GGG GCC CAC CAT GTC GGC TAC ATG TTC CTG GGT GAG GTG GCC CTG 1579
Cys Gly Ala His His Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu
975 980 985

GGC AGA GAG CAC CAT ATC AAC ACG GAC AAC CCC AGC TTG AAG AGC CCA 1627
Gly Arg Glu His His Ile Asn Thr Asp Asn Pro Ser Leu Lys Ser Pro
990 995 1000

CCT CCT GGC TTC GAC AGT GTC ATT GCC CGA GGC CAC ACC GAG CCT GAT 1675
Pro Pro Gly Phe Asp Ser Val Ile Ala Arg Gly His Thr Glu Pro Asp
1005 1010 1015

CCG ACC CAG GAC ACT GAG TTG GAG CTG GAT GGC CAG CAA GTG GTG GTG 1723
Pro Thr Gln Asp Thr Glu Leu Glu Leu Asp Gly Gln Gln Val Val Val
1020 1025 1030 1035

CCC CAG GGC CAG CCT GTG CCC TGC CCA GAG TTC AGC AGC TCC ACA TTC 1771
Pro Gln Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe
1040 1045 1050

TCC CAG AGC GAG TAC CTC ATC TAC CAG GAG AGC CAG TGT CGC CTG CGC 1819
Ser Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg
1055 1060 1065

TAC CTG CTG GAG GTC CAC CTC TGA GTGCCCGCCC TGTCCCCCGG GGTCCTGCAA 1873
Tyr Leu Leu Glu Val His Leu *
1070 1075

GGCTGGACTG TGATCTTCAA TCATCCTGCC CATCTCTGGT ACCCCTATAT CACTCCTTTT 1933

TTTCAAGAAT ACAATACGTT GTTGTTAACT ATAGTCACCA TGCTGTACAA GATCCCTGAA 1993

CTTATGCCTC CTAACTGAAA TTTTGTATTC TTTGACACAT CTGCCCAGTC CCTCTCCTCC 2053

CAGCCCATGG TAACCAGCAT TTGACTCTTT ACTTGTATAA GGGCAGCTTT TATAGGTTCC 2113

ACATGTAAGT GAGATCATGC AGTGTTTGTC TTTCTGTGCC TGGCTTATTT CACTCAGCAT 2173

AATGTGCACC GGGTTCACCC ATGTTTTCAT AAATGACAAG ATTTCCTCCT TTAAAAAAAA 2233

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA 2265

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 541 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Ser Leu Leu Phe Leu Ala Met Ala Pro Lys Pro Lys Pro Trp Val
1 5 10 15

Gln Thr Glu Gly Pro Glu Lys Lys Lys Gly Arg Gln Ala Gly Arg Glu
20 25 30

Glu Asp Pro Phe Arg Ser Thr Ala Glu Ala Leu Lys Ala Ile Pro Ala
35 40 45

Glu Lys Arg Ile Ile Arg Val Asp Pro Thr Cys Pro Leu Ser Ser Asn
50 55 60

Pro Gly Thr Gln Val Tyr Glu Asp Tyr Asn Cys Thr Leu Asn Gln Thr
65 70 75 80

Asn Ile Glu Asn Asn Asn Asn Lys Phe Tyr Ile Ile Gln Leu Leu Gln
85 90 95

Asp Ser Asn Arg Phe Phe Thr Cys Trp Asn Arg Trp Gly Arg Val Gly
100 105 110

Glu Val Gly Gln Ser Lys Ile Asn His Phe Thr Arg Leu Glu Asp Ala
115 120 125

Lys Lys Asp Phe Glu Lys Lys Phe Arg Glu Lys Thr Lys Asn Asn Trp
130 135 140

Ala Glu Arg Asp His Phe Val Ser His Pro Gly Lys Tyr Thr Leu Ile
145 150 155 160

Glu Val Gln Ala Glu Asp Glu Ala Gln Glu Ala Val Val Lys Val Asp
165 170 175

Arg Gly Pro Val Arg Thr Val Thr Lys Arg Val Gln Pro Cys Ser Leu
180 185 190

Asp Pro Ala Thr Gln Lys Leu Ile Thr Asn Ile Phe Ser Lys Glu Met
195 200 205

Phe Lys Asn Thr Met Ala Leu Met Asp Leu Asp Val Lys Lys Met Pro
    210   215   220

Leu Gly Lys Leu Ser Lys Gln Gln Ile Ala Arg Gly Phe Glu Ala Leu
225   230   235   240

Glu Ala Leu Glu Glu Ala Leu Lys Gly Pro Thr Asp Gly Gly Gln Ser
                245   250   255

Leu Glu Glu Leu Ser Ser His Phe Tyr Thr Val Ile Pro His Asn Phe
            260   265   270

Gly His Ser Gln Pro Pro Pro Ile Asn Ser Pro Glu Leu Leu Gln Ala
        275   280   285

Lys Lys Asp Met Leu Leu Val Leu Ala Asp Ile Glu Leu Ala Gln Ala
    290   295   300

Leu Gln Ala Val Ser Glu Gln Glu Lys Thr Val Glu Glu Val Pro His
305   310   315   320

Pro Leu Asp Arg Asp Tyr Gln Leu Leu Lys Cys Gln Leu Gln Leu Leu
                325   330   335

Asp Ser Gly Ala Pro Glu Tyr Lys Val Ile Gln Thr Tyr Leu Glu Gln
            340   345   350

Thr Gly Ser Asn His Arg Cys Pro Thr Leu Gln His Ile Trp Lys Val
        355   360   365

Asn Gln Glu Gly Glu Glu Asp Arg Phe Gln Ala His Ser Lys Leu Gly
    370   375   380

Asn Arg Lys Leu Leu Trp His Gly Thr Asn Met Ala Val Val Ala Ala
385   390   395   400

Ile Leu Thr Ser Gly Leu Arg Ile Met Pro His Ser Gly Gly Arg Val
                405   410   415

Gly Lys Gly Ile Tyr Phe Ala Ser Glu Asn Ser Lys Ser Ala Gly Tyr
            420   425   430

Val Ile Gly Met Lys Cys Gly Ala His His Val Gly Tyr Met Phe Leu
        435   440   445

Gly Glu Val Ala Leu Gly Arg Glu His His Ile Asn Thr Asp Asn Pro

```
    450                 455                 460

Ser Leu Lys Ser Pro Pro Pro Gly Phe Asp Ser Val Ile Ala Arg Gly
465                 470                 475                 480

His Thr Glu Pro Asp Pro Thr Gln Asp Thr Glu Leu Glu Leu Asp Gly
                485                 490                 495

Gln Gln Val Val Val Pro Gln Gly Gln Pro Val Pro Cys Pro Glu Phe
            500                 505                 510

Ser Ser Ser Thr Phe Ser Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser
        515                 520                 525

Gln Cys Arg Leu Arg Tyr Leu Leu Glu Val His Leu  *
    530                 535                 540
```

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1740 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Mus musculus

(ix) FEATURE:
- (A) NAME/KEY: CDS
- (B) LOCATION:112..1710

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
CCCGGCTTTC ACTTTTTCTG CTGCCTCGGG GAACACCTCG AGCCAACTGC TTCCTAACTC       60

AGGGTGGGCA GAACTGACGG GATCTAAGCT TCTGCATCTC TGAGGAGAAC C ATG GCT      117
                                                         Met Ala

CCA AAA CGA AAG GCC TCT GTG CAG ACT GAG GGC TCC AAG AAG CAG CGA      165
Pro Lys Arg Lys Ala Ser Val Gln Thr Glu Gly Ser Lys Lys Gln Arg
    545                 550                 555
```

CAA GGG ACA GAG GAG GAG GAC AGC TTC CGG TCC ACT GCC GAG GCT CTC 213
Gln Gly Thr Glu Glu Glu Asp Ser Phe Arg Ser Thr Ala Glu Ala Leu
560 565 570 575

AGA GCA GCA CCT GCT GAT AAT CGG GTC ATC CGT GTG GAC CCC TCA TGT 261
Arg Ala Ala Pro Ala Asp Asn Arg Val Ile Arg Val Asp Pro Ser Cys
580 585 590

CCA TTC AGC CGG AAC CCC GGG ATA CAG GTC CAC GAG GAC TAT GAC TGT 309
Pro Phe Ser Arg Asn Pro Gly Ile Gln Val His Glu Asp Tyr Asp Cys
595 600 605

ACC CTG AAC CAG ACC AAC ATC GGC AAC AAC AAC AAC AAG TTC TAT ATT 357
Thr Leu Asn Gln Thr Asn Ile Gly Asn Asn Asn Asn Lys Phe Tyr Ile
610 615 620

ATC CAA CTG CTG GAG GAG GGT AGT CGC TTC TTC TGC TGG AAT CGC TGG 405
Ile Gln Leu Leu Glu Glu Gly Ser Arg Phe Phe Cys Trp Asn Arg Trp
625 630 635

GGC CGC GTG GGA GAG GTG GGC CAG AGC AAG ATG AAC CAC TTC ACC TGC 453
Gly Arg Val Gly Glu Val Gly Gln Ser Lys Met Asn His Phe Thr Cys
640 645 650 655

CTG GAA GAT GCA AAG AAG GAC TTT AAG AAG AAA TTT TGG GAG AAG ACT 501
Leu Glu Asp Ala Lys Lys Asp Phe Lys Lys Lys Phe Trp Glu Lys Thr
660 665 670

AAA AAC AAA TGG GAG GAG CGG GAC CGT TTT GTG GCC CAG CCC AAC AAG 549
Lys Asn Lys Trp Glu Glu Arg Asp Arg Phe Val Ala Gln Pro Asn Lys
675 680 685

TAC ACA CTT ATA GAA GTC CAG GGA GAA GCA GAG AGC CAA GAG GCT GTA 597
Tyr Thr Leu Ile Glu Val Gln Gly Glu Ala Glu Ser Gln Glu Ala Val
690 695 700

GTG AAG GCC TTA TCT CCC CAG GTG GAC AGC GGC CCT GTG AGG ACC GTG 645
Val Lys Ala Leu Ser Pro Gln Val Asp Ser Gly Pro Val Arg Thr Val
705 710 715

GTC AAG CCC TGC TCC CTA GAC CCT GCC ACC CAG AAC CTT ATC ACC AAC 693
Val Lys Pro Cys Ser Leu Asp Pro Ala Thr Gln Asn Leu Ile Thr Asn
720 725 730 735

ATC TTC AGC AAA GAG ATG TTC AAG AAC GCA ATG ACC CTC ATG AAC CTG 741
Ile Phe Ser Lys Glu Met Phe Lys Asn Ala Met Thr Leu Met Asn Leu
740 745 750

GAT GTG AAG AAG ATG CCC TTG GGA AAG CTG ACC AAG CAG CAG ATT GCC 789
Asp Val Lys Lys Met Pro Leu Gly Lys Leu Thr Lys Gln Gln Ile Ala
755 760 765

CGT GGC TTC GAG GCC TTG GAA GCT CTA GAG GAG GCC ATG AAA AAC CCC 837
Arg Gly Phe Glu Ala Leu Glu Ala Leu Glu Glu Ala Met Lys Asn Pro
770 775 780

ACA GGG GAT GGC CAG AGC CTG GAA GAG CTC TCC TCC TGC TTC TAC ACT 885
Thr Gly Asp Gly Gln Ser Leu Glu Glu Leu Ser Ser Cys Phe Tyr Thr
785 790 795

GTC ATC CCA CAC AAC TTC GGC CGC AGC CGA CCC CCG CCC ATC AAC TCC 933
Val Ile Pro His Asn Phe Gly Arg Ser Arg Pro Pr Pro Il Asn Ser
800 805 810 815

CCT GAT GTG CTT CAG GCC AAG AAG GAC ATG CTG CTG GTG CTA GCG GAC 981
Pro Asp Val Leu ln Ala Lys Lys Asp Met Leu Leu Val Leu Ala Asp
820 825 830

ATC GAG TTG GCG CAG ACC TTG CAG GCA GCC CCT GGG GAG GAG GAG GAG 1029
Ile Glu Leu Ala Gln Thr Leu Gln Ala Ala Pro Gly Glu Glu Glu Glu
835 840 845

AAA GTG GAA GAG GTG CCA CAC CCA CTG GAT CGA GAC TAC CAG CTC CTC 1077
Lys Val Glu Glu Val Pro His Pro Leu Asp Arg Asp Tyr Gln Leu Leu
850 855 860

AGG TGC CAG CTT CAA CTG CTG GAC TCC GGG GAG TCC GAG TAC AAG GCA 1125
Arg Cys Gln Leu Gln Leu Leu Asp Ser Gly Glu Ser Glu Tyr Lys Ala
865 870 875

ATA CAG ACC TAC CTG AAA CAG ACT GGC AAC AGC TAC AGG TGC CCA AAC 1173
Ile Gln Thr Tyr Leu Lys Gln Thr Gly Asn Ser Tyr Arg Cys Pro Asn
880 885 890 895

CTG CGG CAT GTT TGG AAA GTG AAC CGA GAA GGG GAG GGA GAC AGG TTC 1221
Leu Arg His Val Trp Lys Val Asn Arg Glu Gly Glu Gly Asp Arg Phe
900 905 910

CAG GCC CAC TCC AAA CTG GGC AAT CGG AGG CTG CTG TGG CAC GGC ACC 1269
Gln Ala His Ser Lys Leu Gly Asn Arg Arg Leu Leu Trp His Gly Thr
915 920 925

AAT GTG GCC GTG GTG GCT GCC ATC CTC ACC AGT GGG CTC CGA ATC ATG 1317
Asn Val Ala Val Val Ala Ala Ile Leu Thr Ser Gly Leu Arg Ile Met
930 935 940

CCA CAC TCG GGT GGT CGT GTT GGC AAG GGT ATT TAT TTT GCC TCT GAG 1365

Pro His Ser Gly Gly Arg Val Gly Lys Gly Ile Tyr Phe Ala Ser Glu
945 950 955

AAC AGC AAG TCA GCT GGC TAT GTT ACC ACC ATG CAC TGT GGG GGC CAC 1413
Asn Ser Lys Ser Ala Gly Tyr Val Thr Thr Met His Cys Gly Gly His
960 965 970 975

CAG GTG GGC TAC ATG TTC CTG GGC GAG GTG GCC CTC GGC AAA GAG CAC 1461
Gln Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu Gly Lys Glu His
980 985 990

CAC ATC ACC ATC GAT GAC CCC AGC TTG AAG AGT CCA CCC CCT GGC TTT 1509
His Ile Thr Ile Asp Asp Pro Ser Leu Lys Ser Pro Pro Pro Gly Phe
995 1000 1005

GAC AGC GTC ATC GCC CGA GGC CAA ACC GAG CCG GAT CCC GCC CAG GAC 1557
Asp Ser Val Ile Ala Arg Gly Gln Thr Glu Pro Asp Pro Ala Gln Asp
1010 1015 1020

ATT GAA CTT GAA CTG GAT GGG CAG CCG GTG GTG GTG CCC CAA GGC CCG 1605
Ile Glu Leu Glu Leu Asp Gly Gln Pro Val Val Val Pro Gln Gly Pro
1025 1030 1035

CCT GTG CAG TGC CCG TCA TTC AAA AGC TCC AGC TTC AGC CAG AGT GAA 1653
Pro Val Gln Cys Pro Ser Phe Lys Ser Ser Ser Phe Ser Gln Ser Glu
1040 1045 1050 1055

TAC CTC ATA TAC AAG GAG AGC CAG TGT CGC CTG CGC TAC CTG CTG GAG 1701
Tyr Leu Ile Tyr Lys Glu Ser Gln Cys Arg Leu Arg Tyr Leu Leu Glu
1060 1065 1070

ATT CAC CTC TAAGCTGCTT GCCCTCCCTA GGTCCAAGCC 1740
Il His Leu

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 533 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```
Met Ala Pro Lys Arg Lys Ala Ser Val Gln Thr Glu Gly Ser Lys Lys
1               5                   10                  15

Gln Arg Gln Gly Thr Glu Glu Glu Asp Ser Phe Arg Ser Thr Ala Glu
            20                  25                  30

Ala Leu Arg Ala Ala Pro Ala Asp Asn Arg Val Ile Arg Val Asp Pro
        35                  40                  45

Ser Cys Pro Phe Ser Arg Asn Pro Gly Ile Gln Val His Glu Asp Tyr
    50                  55                  60

Asp Cys Thr Leu Asn Gln Thr Asn Ile Gly Asn Asn Asn Asn Lys Phe
65                  70                  75                  80

Tyr Ile Ile Gln Leu Leu Glu Glu Gly Ser Arg Phe Phe Cys Trp Asn
                85                  90                  95

Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys Met Asn His Phe
            100                 105                 110

Thr Cys Leu Glu Asp Ala Lys Lys Asp Phe Lys Lys Lys Phe Trp Glu
        115                 120                 125

Lys Thr Lys Asn Lys Trp Glu Glu Arg Asp Arg Phe Val Ala Gln Pro
    130                 135                 140

Asn Lys Tyr Thr Leu Ile Glu Val Gln Gly Glu Ala Glu Ser Gln Glu
145                 150                 155                 160

Ala Val Val Lys Ala Leu Ser Pro Gln Val Asp Ser Gly Pro Val Arg
                165                 170                 175

Thr Val Val Lys Pro Cys Ser Leu Asp Pro Ala Thr Gln Asn Leu Ile
            180                 185                 190

Thr Asn Ile Phe Ser Lys Glu Met Phe Lys Asn Ala Met Thr Leu Met
        195                 200                 205

Asn Leu Asp Val Lys Lys Met Pro Leu Gly Lys Leu Thr Lys Gln Gln
    210                 215                 220

Ile Ala Arg Gly Phe Glu Ala Leu Glu Ala Leu Glu Glu Ala Met Lys
225                 230                 235                 240

Asn Pro Thr Gly Asp Gly Gln Ser Leu Glu Glu Leu Ser Ser Cys Phe
                245                 250                 255

Tyr Thr Val Ile Pro His Asn Phe Gly Arg Ser Arg Pro Pro Pro Ile
            260                 265                 270

Asn Ser Pro Asp Val Leu Gln Ala Lys Lys Asp Met Leu Leu Val Leu
        275                 280                 285
```

Ala Asp Ile Glu Leu Ala Gln Thr Leu Gln Ala Ala Pro Gly Glu Glu
290 295 300

Glu Glu Lys Val Glu Glu Val Pro His Pro Leu Asp Arg Asp Tyr Gln
305 310 315 320

Leu Leu Arg Cys Gln Leu Gln Leu Leu Asp Ser Gly Glu Ser Glu Tyr
325 330 335

Lys Ala Ile Gln Thr Tyr Leu Lys Gln Thr Gly Asn Ser Tyr Arg Cys
340 345 350

Pro Asn Leu Arg His Val Trp Lys Val Asn Arg Glu Gly Glu Gly Asp
355 360 365

Arg Phe Gln Ala His Ser Lys Leu Gly Asn Arg Arg Leu Leu Trp His
370 375 380

Gly Thr Asn Val Ala Val Val Ala Ala Ile Leu Thr Ser Gly Leu Arg
385 390 395 400

Ile Met Pro His Ser Gly Gly Arg Val Gly Lys Gly Ile Tyr Phe Ala
405 410 415

Ser Glu Asn Ser Lys Ser Ala Gly Tyr Val Thr Thr Met His Cys Gly
420 425 430

Gly His Gln Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu Gly Lys
435 440 445

Glu His His Ile Thr Ile Asp Asp Pro Ser Leu Lys Ser Pro Pro Pro
450 455 460

Gly Phe Asp Ser Val Ile Ala Arg Gly Gln Thr Glu Pro Asp Pro Ala
465 470 475 480

Gln Asp Ile Glu Leu Glu Leu Asp Gly Gln Pro Val Val Val Pro Gln
485 490 495

Gly Pro Pro Val Gln Cys Pro Ser Phe Lys Ser Ser Ser Phe Ser Gln
500 505 510

Ser Glu Tyr Leu Ile Tyr Lys Glu Ser Gln Cys Arg Leu Arg Tyr Leu
515 520 525

Leu Glu Ile His Leu
530

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1587 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mus musculus

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..1584

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

ATG GCT CCA AAA CGA AAG GCC TCT GTG CAG ACT GAG GGC TCC AAG AAG 48
Met Ala Pro Lys Arg Lys Ala Ser Val Gln Thr Glu Gly Ser Lys Lys
535 540 545

CAG CGA CAA GGG ACA GAG GAG GAG GAC AGC TTC CGG TCC ACT GCC GAG 96
Gln Arg Gln Gly Thr Glu Glu Glu Asp Ser Phe Arg Ser Thr Ala Glu
550 555 560 565

GCT CTC AGA GCA GCA CCT GCT GAT AAT CGG GTC ATC CGT GTG GAC CCC 144
Ala Leu Arg Ala Ala Pro Ala Asp Asn Arg Val Ile Arg Val Asp Pro
570 575 580

TCA TGT CCA TTC AGC CGG AAC CCC GGG ATA CAG GTC CAC GAG GAC TAT 192
Ser Cys Pro Phe Ser Arg Asn Pro Gly Ile Gln Val His Glu Asp Tyr
585 590 595

GAC TGT ACC CTG AAC CAG ACC AAC ATC GGC AAC AAC AAC AAC AAG TTC 240
Asp Cys Thr Leu Asn Gln Thr Asn Ile Gly Asn Asn Asn Asn Lys Phe
600 605 610

TAT ATT ATC CAA CTG CTG GAG GAG GGT AGT CGC TTC TTC TGC TGG AAT 288
Tyr Ile Ile Gln Leu Leu Glu Glu Gly Ser Arg Phe Phe Cys Trp Asn
615 620 625

CGC TGG GGC CGC GTG GGA GAG GTG GGC CAG AGC AAG ATG AAC CAC TTC 336
Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys Met Asn His Phe
630 635 640 645

ACC TGC CTG GAA GAT GCA AAG AAG GAC TTT AAG AAG AAA TTT TGG GAG 384
Thr Cys Leu Glu Asp Ala Lys Lys Asp Phe Lys Lys Lys Phe Trp Glu
650 655 660

AAG ACT AAA AAC AAA TGG GAG GAG CGG GAC CGT TTT GTG GCC CAG CCC 432
Lys Thr Lys Asn Lys Trp Glu Glu Arg Asp Arg Phe Val Ala Gln Pro
665 670 675

AAC AAG TAC ACA CTT ATA GAA GTC CAG GGA GAA GCA GAG AGC CAA GAG 480
Asn Lys Tyr Thr Leu Ile Glu Val Gln Gly Glu Ala Glu Ser Gln Glu
680 685 690

GCT GTA GTG AAG GTG GAC AGC GGC CCT GTG AGG ACC GTG GTC AAG CCC 528
Ala Val Val Lys Val Asp Ser Gly Pro Val Arg Thr Val Val Lys Pro
695 700 705

TGC TCC CTA GAC CCT GCC ACC CAG AAC CTT ATC ACC AAC ATC TTC AGC 576
Cys Ser Leu Asp Pro Ala Thr Gln Asn Leu Ile Thr Asn Ile Phe Ser
710 715 720 725

AAA GAG ATG TTC AAG AAC GCA ATG ACC CTC ATG AAC CTG GAT GTG AAG 624
Lys Glu Met Phe Lys Asn Ala Met Thr Leu Met Asn Leu Asp Val Lys
730 735 740

AAG ATG CCC TTG GGA AAG CTG ACC AAG CAG CAG ATT GCC CGT GGC TTC 672
Lys Met Pro Leu Gly Lys Leu Thr Lys Gln Gln Ile Ala Arg Gly Phe
745 750 755

GAG GCC TTG GAA GCT CTA GAG GAG GCC ATG AAA AAC CCC ACA GGG GAT 720
Glu Ala Leu Glu Ala Leu Glu Glu Ala Met Lys Asn Pro Thr Gly Asp
760 765 770

GGC CAG AGC CTG GAA GAG CTC TCC TCC TGC TTC TAC ACT GTC ATC CCA 768
Gly Gln Ser Leu Glu Glu Leu Ser Ser Cys Phe Tyr Thr Val Ile Pro
775 780 785

CAC AAC TTC GGC CGC AGC CGA CCC CCG CCC ATC AAC TCC CCT GAT GTG 816
His Asn Phe Gly Arg Ser Arg Pro Pro Pro Ile Asn Ser Pro Asp Val
790 795 800 805

CTT CAG GCC AAG AAG GAC ATG CTG CTG GTG CTA GCG GAC ATC GAG TTG 864
Leu Gln Ala Lys Lys Asp Met Leu Leu Val Leu Ala Asp Ile Glu Leu
810 815 820

GCG CAG ACC TTG CAG GCA GCC CCT GGG GAG GAG GAG GAG AAA GTG GAA 912
Ala Gln Thr Leu Gln Ala Ala Pro Gly Glu Glu Glu Glu Lys Val Glu
825 830 835

GAG GTG CCA CAC CCA CTG GAT CGA GAC TAC CAG CTC CTC AGG TGC CAG 960
Glu Val Pro His Pro Leu Asp Arg Asp Tyr Gln Leu Leu Arg Cys Gln
840 845 850

CTT CAA CTG CTG GAC TCC GGG GAG TCC GAG TAC AAG GCA ATA CAG ACC 1008
Leu Gln Leu Leu Asp Ser Gly Glu Ser Glu Tyr Lys Ala Ile Gln Thr
855 860 865

TAC CTG AAA CAG ACT GGC AAC AGC TAC AGG TGC CCA AAC CTG CGG CAT 1056
Tyr Leu Lys Gln Thr Gly Asn Ser Tyr Arg Cys Pro Asn Leu Arg His
870 875 880 885

GTT TGG AAA GTG AAC CGA GAA GGG GAG GGA GAC AGG TTC CAG GCC CAC 1104

Val Trp Lys Val Asn Arg Glu Gly Glu Gly Asp Arg Phe Gln Ala His
890 895 900

TCC AAA CTG GGC AAT CGG AGG CTG CTG TGG CAC GGC ACC AAT GTG GCC 1152
Ser Lys Leu Gly Asn Arg Arg Leu Leu Trp His Gly Thr Asn Val Ala
905 910 915

GTG GTG GCT GCC ATC CTC ACC AGT GGG CTC CGA ATC ATG CCA CAC TCG 1200
Val Val Ala Ala Ile Leu Thr Ser Gly Leu Arg Ile Met Pro His Ser
920 925 930

GGT GGT CGT GTT GGC AAG GGT ATT TAT TTT GCC TCT GAG AAC AGC AAG 1248
Gly Gly Arg Val Gly Lys Gly Ile Tyr Phe Ala Ser Glu Asn Ser Lys
935 940 945

TCA GCT GGC TAT GTT ACC ACC ATG CAC TGT GGG GGC CAC CAG GTG GGC 1296
Ser Ala Gly Tyr Val Thr Thr Met His Cys Gly Gly His Gln Val Gly
950 955 960 965

TAC ATG TTC CTG GGC GAG GTG GCC CTC GGC AAA GAG CAC CAC ATC ACC 1344
Tyr Met Phe Leu Gly Glu Val Ala Leu Gly Lys Glu His His Ile Thr
970 975 980

ATC GAT GAC CCC AGC TTG AAG AGT CCA CCC CCT GGC TTT GAC AGC GTC 1392
Ile Asp Asp Pro Ser Leu Lys Ser Pro Pro Pro Gly Phe Asp Ser Val
985 990 995

ATC GCC CGA GGC CAA ACC GAG CCG GAT CCC GCC CAG GAC ATT GAA CTT 1440
Ile Ala Arg Gly Gln Thr Glu Pro Asp Pro Ala Gln Asp Ile Glu Leu
1000 1005 1010

GAA CTG GAT GGG CAG CCG GTG GTG GTG CCC CAA GGC CCG CCT GTG CAG 1488
Glu Leu Asp Gly Gln Pro Val Val Val Pro Gln Gly Pro Pro Val Gln
1015 1020 1025

TGC CCG TCA TTC AAA AGC TCC AGC TTC AGC CAG AGT GAA TAC CTC ATA 1536
Cys Pro Ser Phe Lys Ser Ser Ser Phe Ser Gln Ser Glu Tyr Leu Ile
1030 1035 1040 1045

TAC AAG GAG AGC CAG TGT CGC CTG CGC TAC CTG CTG GAG ATT CAC CTC 1584
Tyr Lys Glu Ser Gln Cys Arg Leu Arg Tyr Leu Leu Glu Ile His Leu
1050 1055 1060

TAA 1587

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 528 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```
Met Ala Pro Lys Arg Lys Ala Ser Val Gln Thr Glu Gly Ser Lys Lys
1               5                   10                  15

Gln Arg Gln Gly Thr Glu Glu Glu Asp Ser Phe Arg Ser Thr Ala Glu
            20                  25                  30

Ala Leu Arg Ala Ala Pro Ala Asp Asn Arg Val Ile Arg Val Asp Pro
        35                  40                  45

Ser Cys Pro Phe Ser Arg Asn Pro Gly Ile Gln Val His Glu Asp Tyr
    50                  55                  60

Asp Cys Thr Leu Asn Gln Thr Asn Ile Gly Asn Asn Asn Asn Lys Phe
65                  70                  75                  80

Tyr Ile Ile Gln Leu Leu Glu Glu Gly Ser Arg Phe Phe Cys Trp Asn
                85                  90                  95

Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys Met Asn His Phe
            100                 105                 110

Thr Cys Leu Glu Asp Ala Lys Lys Asp Phe Lys Lys Lys Phe Trp Glu
        115                 120                 125

Lys Thr Lys Asn Lys Trp Glu Glu Arg Asp Arg Phe Val Ala Gln Pro
    130                 135                 140

Asn Lys Tyr Thr Leu Ile Glu Val Gln Gly Glu Ala Glu Ser Gln Glu
145                 150                 155                 160

Ala Val Val Lys Val Asp Ser Gly Pro Val Arg Thr Val Val Lys Pro
                165                 170                 175

Cys Ser Leu Asp Pro Ala Thr Gln Asn Leu Ile Thr Asn Ile Phe Ser
            180                 185                 190

Lys Glu Met Phe Lys Asn Ala Met Thr Leu Met Asn Leu Asp Val Lys
        195                 200                 205

Lys Met Pro Leu Gly Lys Leu Thr Lys Gln Gln Ile Ala Arg Gly Phe
    210                 215                 220

Glu Ala Leu Glu Ala Leu Glu Glu Ala Met Lys Asn Pro Thr Gly Asp
225                 230                 235                 240

Gly Gln Ser Leu Glu Glu Leu Ser Ser Cys Phe Tyr Thr Val Ile Pro
                245                 250                 255

His Asn Phe Gly Arg Ser Arg Pro Pr  Pro Ile Asn Ser Pro Asp Val
            260                 265                 270

Leu Gln Ala Lys Lys Asp Met Leu Leu Val Leu Ala Asp Il  Glu Leu
        275                 280                 285
```

Ala Gln Thr Leu Gln Ala Ala Pro Gly Glu Glu Glu Glu Lys Val Glu
290 295 300

Glu Val Pro His Pro Leu Asp Arg Asp Tyr Gln Leu Leu Arg Cys Gln
305 310 315 320

Leu Gln Leu Leu Asp Ser Gly Glu Ser Glu Tyr Lys Ala Ile Gln Thr
325 330 335

Tyr Leu Lys Gln Thr Gly Asn Ser Tyr Arg Cys Pro Asn Leu Arg His
340 345 350

Val Trp Lys Val Asn Arg Glu Gly Glu Gly Asp Arg Phe Gln Ala His
355 360 365

Ser Lys Leu Gly Asn Arg Arg Leu Leu Trp His Gly Thr Asn Val Ala
370 375 380

Val Val Ala Ala Ile Leu Thr Ser Gly Leu Arg Ile Met Pro His Ser
385 390 395 400

Gly Gly Arg Val Gly Lys Gly Ile Tyr Phe Ala Ser Glu Asn Ser Lys
405 410 415

Ser Ala Gly Tyr Val Thr Thr Met His Cys Gly Gly His Gln Val Gly
420 425 430

Tyr Met Phe Leu Gly Glu Val Ala Leu Gly Lys Glu His His Ile Thr
435 440 445

Ile Asp Asp Pro Ser Leu Lys Ser Pro Pro Pro Gly Phe Asp Ser Val
450 455 460

Ile Ala Arg Gly Gln Thr Glu Pro Asp Pro Ala Gln Asp Ile Glu Leu
465 470 475 480

Glu Leu Asp Gly Gln Pro Val Val Val Pro Gln Gly Pro Pro Val Gln
485 490 495

Cys Pro Ser Phe Lys Ser Ser Ser Phe Ser Gln Ser Glu Tyr Leu Ile
500 505 510

Tyr Lys Glu Ser Gln Cys Arg Leu Arg Tyr Leu Leu Glu Ile His Leu
515 520 525

We claim:

1. A poly(ADP-ribose) polymerase (PARP) homolog which has an amino acid sequence which has
   a) a functional NAD^+ binding domain and
   b) no zinc finger sequence motif of the general formula

$$CX_2CX_mHX_2C$$

in which
m is an integral value from 28 or 30, and the X radicals are, independently of one another, any amino acid;
and the functional equivalents thereof.

2. A PARP homolog as claimed in claim 1, wherein the functional NAD^+ binding domain comprises one of the following general sequence motifs:

$PX_n(S/T)GX_3GKGIYFA$,
$(S/T)XGLR(I/V)XPX_n(S/T)GX_3GKGIYFA$ or
$LLWHG(S/T)X_7IL(S/T)XGLR(I/V)XPX_n(S/T)GX_3GKGIYFAX_3SKSAXY$

in which
n is an integral value from 1 to 5, and the X radicals are, independently of one another, any amino acid.

3. A PARP homolog as claimed in either of the preceding claims, comprising at least another one of the following part-sequence motifs:

$LX_9NX_2YX_2QLLX(D/E)X_{10/11}WGRVG$,
$AX_3FXKX_4KTXNXWX_5FX_3PXK$,
$QXL(I/L)X_2IX_9MX_{10}PLGKLX_3QIX_6L$,
$FYTXIPHXFGX_3PP$; and
$KX_3LX_2LXDIEXAX_2L$,

in which the X radicals are, independently of one another, any amino acid.

4. A PARP homolog as claimed in any of the preceding claims, selected from human PARP homologs, which has the amino acid sequence shown in SEQ ID NO: 2 (human PARP2) or SEQ ID NO: 4 or 6 (human PARP3 type 1 or 2); or murine PARP homologs which have the amino acid sequence shown in SEQ ID NO:8 (mouse PARP

long form) or SEQ ID No:10 (mouse PARP short form); and the functional equivalents thereof.

5. A binding partner for PARP homologs as claimed in any of the preceding claims, selected from
   a) antibodies and fragments thereof,
   b) protein-like compounds which interact with a part-sequence of the protein, and
   c) low molecular weight effectors which modulate the catalytic PARP activity or another biological function of a PARP molecule.

6. A nucleic acid comprising
   a) a nucleotide sequence coding for at least one PARP homolog as claimed in any of claims 1 to 4, or the complementary nucleotide sequence thereof;
   b) a nucleotide sequence which hybridizes with a sequence as specified in a) under stringent conditions; or
   c) nucleotide sequences which are derived from the nucleotide sequences defined in a) and b) through the degeneracy of the genetic code.

7. A nucleic acid as claimed in claim 6, comprising

   a) nucleotides +3 to +1715 shown in SEQ ID NO:1;
   b) nucleotides +242 to +1843 shown in SEQ ID NO:3;
   c) nucleotides +221 to +1843 shown in SEQ ID NO:5;
   d) nucleotides +112 to +1710 shown in SEQ ID NO:7; or
   e) nucleotides +1 to +1584 shown in SEQ ID NO:9.

8. An expression cassette comprising, under the genetic control of at least one regulatory nucleotide sequence, at least one nucleotide sequence as claimed in either of claims 6 and 7.

9. A recombinant vector comprising at least one expression cassette as claimed in claim 8.

10. A recombinant microorganism comprising at least one recombinant vector as claimed in claim 9.

11. A transgenic mammal comprising a vector as claimed in claim 9.

12. A PARP-deficient mammal or PARP-deficient eukaryotic cell, in which functional expression of at least one gene which codes for a PARP homolog as claimed in any of claims 1 to 4 is inhibited.

13. An in vitro detection method for PARP inhibitors, which comprises

a) incubating an unsupported or supported polyADP-ribosylatable target with a reaction mixture comprising
   a1) a PARP homolog as claimed in any of claims 1 to 4,
   a2) a PARP activator; and
   a3) a PARP inhibitor or an analyte in which at least one PARP inhibitor is suspected;
b) carrying out the polyADP ribosylation reaction; and
c) determining the polyADP ribosylation of the target qualitatively or quantitatively.

14. A method as claimed in claim 13, wherein the PARP homolog is preincubated with the PARP activator and the PARP inhibitor or an analyte in which at least one PARP inhibitor is suspected, before the polyADP ribosylation reaction is carried out.

15. A method as claimed in either of claims 13 and 14, wherein the polyADP-ribosylatable target is a histone protein.

16. A method as claimed in any of claims 13 to 15, wherein the PARP activator is activated DNA.

17. A method as claimed in any of claims 13 to 16, wherein the polyADP ribosylation reaction is started by adding NAD^+.

18. A method as claimed in any of claims 13 to 17, wherein the polyADP ribosylation of the supported target is determined using anti-poly(ADP-ribose) antibodie.

19. A method as claimed in any of claims 13 to 17, wherein the unsupported target is labeled with an acceptor fluorophore.

20. A method as claimed in claim 19, wherein the polyADP ribosylation of the unsupported target is determined using anti-poly(ADP-ribose) antibody which is labeled with a donor fluorophore which is able to transfer energy to the acceptor fluorophore.

21. A method as claimed in either of claims 19 and 20, wherein the target is biotinylated histone, and the acceptor fluorophore is coupled thereto via avidin or streptavidin.

22. A method as claimed in either of claims 20 and 21, wherein the anti-poly(ADP-ribose) antibody carries a europium cryptate as donor fluorophore.

23. An in vitro screening method for binding partners for a PARP molecule, which comprises
   a1) immobilizing at least one PARP homolog as claimed in any of claims 1 to 4 on a support;
   b1) contacting the immobilized PARP homolog with an analyte in which at least one binding partner is suspected; and
   c1) determining, where appropriate after an incubation period, analyte constituents bound to the immobilized PARP homolog;

   or

   a2) immobilizing on a support an analyte which comprises at least one possible binding partner for a PARP molecule;
   b2) contacting the immobilized analyte with at least one PARP homolog as claimed in any of claims 1 to 4 for which a binding partner is sought; and
   c2) examining the immobilized analyte, where appropriate after an incubation period, for binding of the PARP homolog.

24. A method for the qualitative or quantitative determination of nucleic acids encoding a PARP homolog as claimed in any of claims 1 to 4, which comprises
   a) incubating a biological sample with a defined amount of an exogenous nucleic acid as claimed in either of claims 6 and 7, hybridizing under stringent conditions, determining the hybridizing nucleic acids and, where appropriate, comparing with a standard; or

   b) incubating a biological sample with a pair of oligonucleotide primers with specificity for a PARP homolog-encoding nucleic acid, amplifying the nucleic acid, determining the amplification product and, where appropriate, comparing with a standard.

25. A method for the qualitative or quantitative determination of a PARP homolog as claimed in any of claims 1 to 4, which comprises

a) incubating a biological sample with a binding partner specific for a PARP homolog,
b) detecting the binding partner/PARP complex and, where appropriate,
c) comparing the result with a standard.

26. A method as claimed in claim 25, wherein the binding partner is an antibody or a binding fragment thereof, which carries a detectable label where appropriate.

27. A method as claimed in any of claims 24 to 26 for diagnosing energy deficit-mediated illnesses.

28. A method for determining the efficacy of PARP effectors, which comprises
a) incubating a PARP homolog as claimed in any of claims 1 to 4 with an analyte which comprises an effector of a physiological or pathological PARP activity; removing the effector again where appropriate; and
b) determining the activity of the PARP homolog, where appropriate after adding substrates or cosubstrates.

29. A gene therapy composition, which comprises in a vehicle acceptable for gene therapy a nucleic acid construct which
a) comprises an antisense nucleic acid against a coding nucleic acid as claimed in either of claims 6 and 7; or
b) a ribozyme against a nucleic acid as claimed in either of claims 6 and 7; or
c) codes for a specific PARP inhibitor.

30. A pharmaceutical composition comprising, in a pharmaceutically acceptable vehicle, at least one PARP protein as claimed in any of claims 1 to 4, at least one PARP binding partner as claimed in claim 5 or at least one coding nucleotide sequence as claimed in claim 6 or 7.

31. The use of low molecular weight PARP binding partners as claimed in claim 5 for the diagnosis or therapy of pathological states in the development and/or progress of which at least one PARP protein, or a polypeptide derived therefrom, are [sic] involved.

32. The use of low molecular weight PARP binding partners as claimed in claim 5 for the diagnosis or therapy of pathological states mediated by an energy deficit.

58/iT/cb

Abstract

The invention relates to poly(ADP-ribose) polymerase (PARP) homologs which have an amino acid sequence which has

a) a functional NAD^+ binding domain
and

b) no zinc finger sequence motif of the general formula

$$CX_2CX_mHX_2C$$

in which
m is an integral value from 28 or 30, and the X radicals are, independently of one another, any amino acid;

and the functional equivalents thereof; nucleic acids coding therefor; antibodies with specificity for the novel protein; pharmaceutical and gene therapy compositions which comprise products according to the invention; methods for the analytical determination of the proteins and nucleic acids according to the invention; methods for identifying effectors or binding partners of the proteins according to the invention; novel PARP effectors; and methods for determining the activity of such effectors.

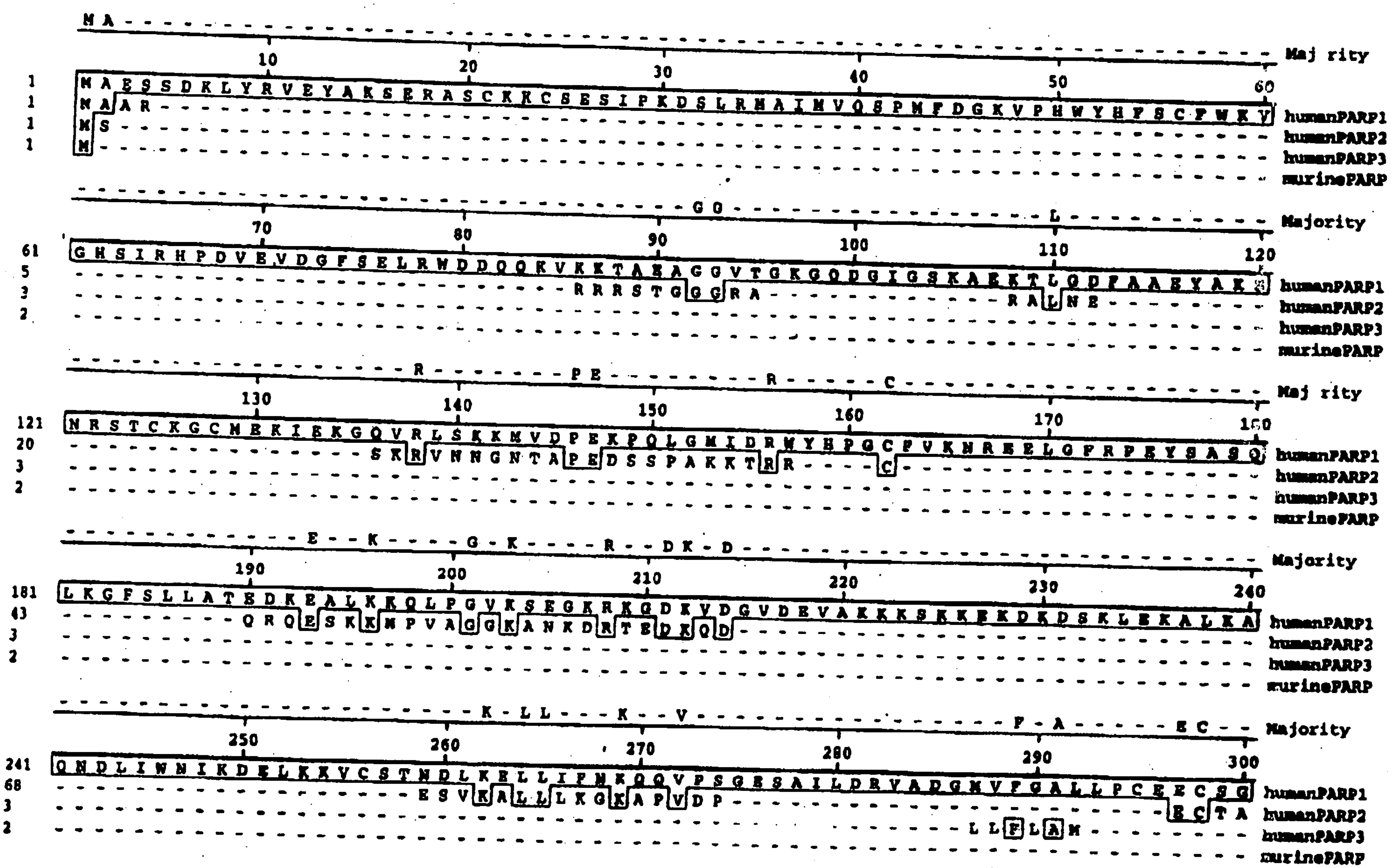

Majority
humanPARP1
humanPARP2
humanPARP3
murinePARP
MAESSDKLYRVEYAKSGRASCKKCSESIPKDSLRMAIMVQSPMFDGKVPHWYHFSCFWKV
GHSIRHPDVEVDGFSELRWDDQQKVKKTAEAGGVTGKGQDGIGSKAEKTLGDFAAEYAKS
NRSTCKGCMEKIEKGQVRLSKKMVDPEKPQLGMIDRWYHPGCFVKNREELGFRPEYSASQ
LKGFSLLATEDKEALKKQLPGVKSEGKRKGDEVDGVDEVAKKKSKKEKDKDSKLEKALKA
QNDLIWNIKDELKKVCSTNDLKELLIFNKQQVPSGESAILDRVADGMVFGALLPCEECSG

Fig. 1(1)

```
                   310        320        330        340        350        360
    ----------YC-G---------------APKRKXWV---------------------Q  Majority
301 QLVFKSDAYYCTGDVTAWTKCMVKTQTPNRKEWVTPKEFREISYLKKLKVKKQDRIFPPE  humanPARP1
88  KVG--KAHVYCEGN----------------------------------------------  humanPARP2
9   ----------------------------APKPKPWV-----------------------Q  humanPARP3
2   ----------------------------APKRKASV-----------------------Q  murinePARP

                   370        380        390        400        410        420
    TEGS--------------------------------------------------------  Maj rity
361 TSASVAATPPPSTASAPAAVNSSASADKPLSNMKILTLGKLSRNKDEVKAMIEKLGGKLT  humanPARP1
100 ------------------------------------------------------------  humanPARP2
18  TEGP--------------------------------------------------------  humanPARP3
11  TEGS--------------------------------------------------------  murinePARP

                   430        440        450        460        470        480
    ---------------EKKKXRQXXXEEDXFRSTAEAL-----------------------  Majority
421 GTANKASLCISTKKEVEKMNKKMEEVKEANIRVVSEDFLQDVSASTKSLQELFLAHILSP  humanPARP1
100 ------------------------------------------------------------  humanPARP2
22  ---------------EKKKGRQAGREEDPPRSTAEAL-----------------------  humanPARP3
15  -----------------KKQRQGTEEEDSFRSTAEAL-----------------------  murinePARP

                   490        500        510        520        530        540
    -----KAXPAEXRXIRVDPXCPLSXNPGXQVXED--------------------------  Maj rity
481 WGAEVKAEPVEVVAPRGKSGAALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLE  humanPARP1
100 ------------------------------------------------------------  humanPARP2
44  -----KAIPAEKRIIRVDPTCPLSSNPGTQVYED--------------------------  humanPARP3
35  -----RAAPADNRVIRVDPSCPFSRNPGIQVHED--------------------------  murinePARP

                   550        560        570        580        590        600
    -----------VYDCTLNQTNIXNNNNKFYIIQLLEDDX-RFFXCWNRWGRVGE-VGQSK  Maj rity
541 HSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSNK  humanPARP1
100 ----------DVYDVMLNQTNLQPNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKM-GQHS  humanPARP2
73  -----------YNCTLNQTNIENNNNKFYIIQLLQDSN-RFFTCWNRWGRVGE-VGQSK  humanPARP3
64  -----------YDCTLNQTNIGNNNHKRFYIIQLLEEGS-RFF-CWNRWGRVGE-VGQSK  murinePARP
```

Fig. 1(2)

```
     L N H P T X - L E D A K E D F X K K F X E R T K N N W E E R D X F V K X P G K Y T L L E V D Y - X E X E D E E A V V K -   Majority
                 610                 620                 630                 640                 650                 660
601  L E Q M P S K - E D A I E H P M K L Y E E K T G N A W H S K N - F T K Y P K K F Y P L E I D Y G - - - Q D E E A V K K -   humanPARP1
149  L V A C S G N L N K A K E I P Q K K F L D R T K N N W E D R E K F E K V P G K Y D N L Q N D Y A T N T Q D E E E T K K E   humanPARP2
119  I N H P T R - L E D A K K D F E K K K F R E K T K N N W A E R D H F V S H P G K Y T L I E V Q - - A E D E A Q E A V V K -   humanPARP3
109  M N H P T C - L E D A K K D F K K K K P W E K T K N K W E E R D R F V A Q P N K Y T L I E V Q - - G E A E S Q E A V V K A   murinePARP

     - S L X V D X G P V S T V X K R V Q P C S L D P A T Q X L I T N I F S V E M P K N A M X L M X L D V K K M P L G K L S K   Maj rity
                 670                 680                 690                 700                 710                 720
655  - - L T V N P G T K S K L P K P V Q - - - - - - - - - - D L I K M I P D V E S M K K A N V E Y E I D L Q K M P L G K L S K   humanPARP1
209  E S L K S P L K P E S Q L D L R V Q - - - - - - - - - - E L I K L I C N V Q A M E E M M M E M K Y N T K P A P L G K L T V   humanPARP2
175  - - - - V D R G P V R T V T K R V Q P C S L D P A T Q K L I T N I F S K E M F K N T M A L M D L D V K K M P L G K L S K   humanPARP3
166  L S P Q V D S G P V R T V V K - - - P C S L D P A T Q N L I T N I F S K E M F K N A M T L M N L D V K K M P L G K L T E   murinePARP

     Q Q I A A G F E A L E A L E E A X K X G T X G G Q S L E E L S S X F Y T V I P H D F G X S X P P L I N S P D X L Q A K X   Maj rity
                 730                 740                 750                 760                 770                 780
704  R Q I Q A A Y S I L S E V Q Q A V S Q G S S D S Q I L D - L S N R P Y T L I P H D F G M K K P P L L N N A D S V Q A K V   humanPARP1
260  A Q I K A G Y Q S L K K I E D C I R A G Q H G R A L M E - A C N E F Y T R I P H D F G L R T P P L I R T Q K E L S E K I   humanPARP2
231  Q Q I A R G F E A L E A L E E A L K G P T D G G Q S L E E L S S H F Y T V I P H N F G H S Q P P P I N S P E L L Q A K K   humanPARP3
223  Q Q I A R G F E A L E A L E E A M K N P T G D G Q S L E E L S S C F Y T V I P H N F G R S R P P F I N S P D V L Q A K K   murinePARP

     D M L L V L A D I E L A Q X L Q A X X X E X S X K V E E V P H P L D R D Y Q L L K C Q L Q L L D S G S X E Y K V I Q T Y   Maj rity
                 790                 800                 810                 820                 830                 840
763  E M L D N L L D I E V A Y S L L R G G S D D S S K - - - - - D P I D V N Y E K L K T D I K V V D R D S E E A E I I R K Y   humanPARP1
319  Q L L E A L G D I E I A I K L V K T E L Q - S P E - - - - - - H P L D Q H Y R N L H C A L R P L D H E S Y E F K V I S Q Y   humanPARP2
291  D M L L V L A D I E L A Q A L Q A V S - E Q E K T V E E V P H P L D R D Y Q L L K C Q L Q L L D S G A P E Y K V I Q T Y   humanPARP3
283  D M L L V L A D I E L A Q T L Q A A P G E E E E K V E E V P H P L D R D Y Q L L R C Q L Q L L D S G E S E Y K A I Q T Y   murinePARP

     L K Q T G A X T H C P Y - - T L X D I P K V E R E G E X D R F Q A H S K L G N R R L L W H G S N M A V V A G I L S S G L   Majority
                 850                 860                 870                 880                 890                 900
818  V K N T H A T T H N A Y D L E V I D I P K I E R E G E C Q R Y K P F K Q L H N R R L L W H G S R T T N F A G I L S Q G L   humanPARP1
373  L Q S T H A P T H S D Y T M T L L D L F E V E K D G E K E A P R - - E D L H N R M L L W H G S R M S N W V G I L S H G L   humanPARP2
350  L E Q T G S N H R C P - - - T L Q H I W K V N Q E G E E D R F Q A H S K L G N R K L L W H G T N M A V V A A I L T S G L   humanPARP3
343  L K Q T G N S Y R C P - - - N L R H V W K V N R E G E G D K F Q A H S K L G N R R L L W H G T N V A V V A A I L T S G L   murinePARP
```

Fig. 1(3)

```
                                  910                 920                 930                 940                 950                 960
    R I A P H E A P - S G G R V G K G I Y P A S E N S K S A G Y V X T S X C G G H X V G L H L L G E V A L G X E H E L X X A  Majority
878 R I A P P E A P V T G Y N F G K G I Y F A D N V S K S A N Y C H T S Q - - G D P I G L I L L G E V A L G N M Y E L K H A  humanPARP1
431 R I A P P E A P I T G Y M P G K G I Y F A D M S S K S A N Y C F A S R - - L K N T G L L L L S E V A L G Q C N E L L E A  humanPARP2
407 R I N P H - - - - S G G R V G K G I Y F A S E N S K S A G Y V I G M K C G A H H V G Y M F L G E V A L G R E H H I N T S  humanPARP3
400 R I N P H - - - - S G G R V G K G I Y F A S E N S K S A G Y V T T M H C G G H Q V G Y N P L G E V A L G K E H H I T I D  murinePARP

                                  970                 980                 990                 1000                1010                1020
    N P S L K S L P P G K D S V I G L G K T E P D P A Q D I E L E L D G Q G V V V P L G P P V X C G X P X S S X F S L - Y G  Majority
936 S H I S K - L P K G K H S V K G L G K T T P D P S A N I S L D - - - - G V D V P L G T G I S S G V - - - N D T S L L Y N  humanPARP1
489 N P K A E G L L Q G K H S T K G L G K M A P S S A H F V T L N - - - - G S T V P L G P A S D T G I L N P D G Y T L N Y N  humanPARP2
463 N P S L K S P P P G F D S V I A R G H T E P D P T Q D T E L E L D G Q Q V V V P Q G Q P V P C P E F S S S T F S Q - - S  humanPARP3
456 D P S L K S P P P G F D S V I A R G Q T E P D P A Q D I E L E L D G Q P V V V P Q G P P V Q C P S F K S S S F S Q - - S  murinePARP

                          1030                1040
    E Y L V Y X E S Q V R L R Y L L E V H F N F - X X L W -  Majority
988 E Y I V Y D I A Q V N L K Y L L K L K F N F K T S L W .  humanPARP1
545 E Y I V Y N P N Q V R M R Y L L K V Q F N F - L Q L W .  humanPARP2
521 E Y L I Y Q E S Q C R L R Y L L E V H L .                humanPARP3
514 E Y L I Y K E S Q C R L R Y L L E I H - - - - - - - - - L  murinePARP
```

Fig. 1(4)

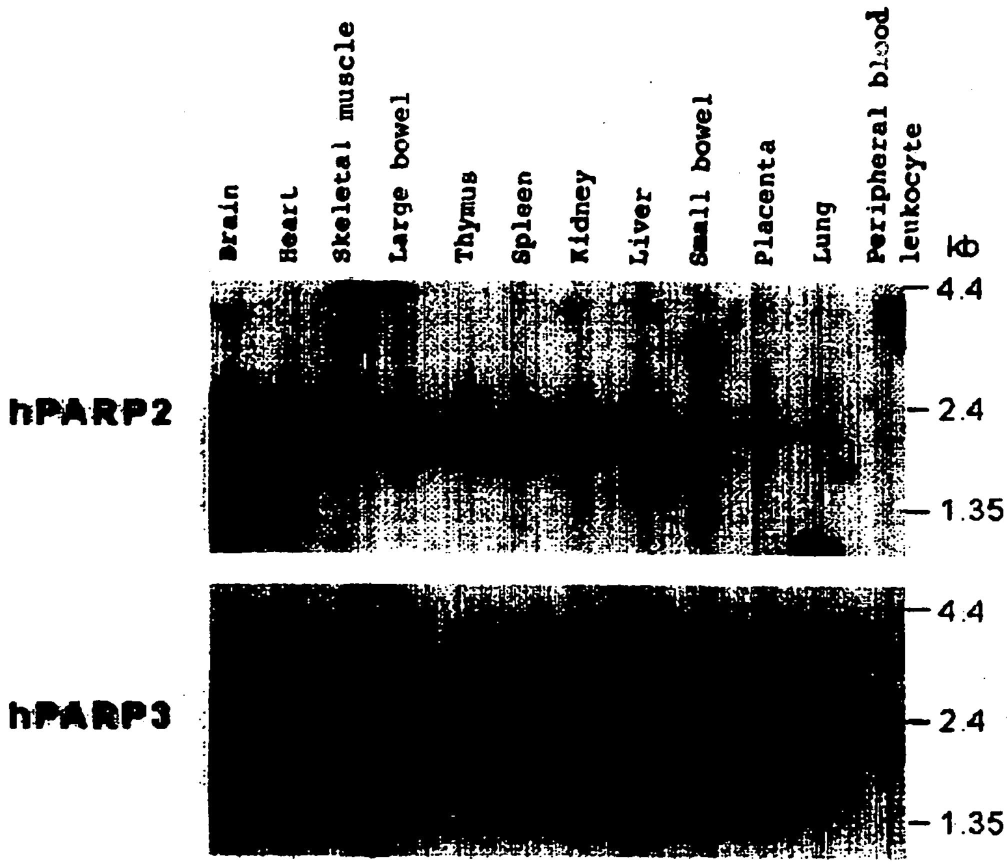

Brain
Heart
Skeletal muscle
Large bowel
Thymus
Spleen
Kidney
Liver
Small bowel
Placenta
Lung
Peripheral blood leukocyte
kb
4.4
2.4
1.35
hPARP2
4.4
2.4
1.35
hPARP3

Fig. 2

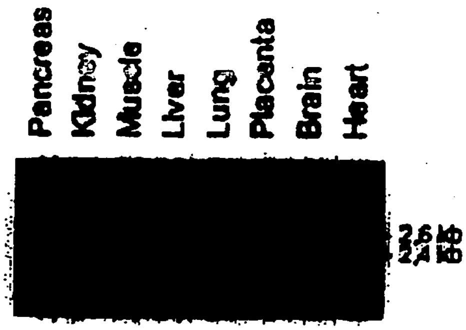


Fig. 3

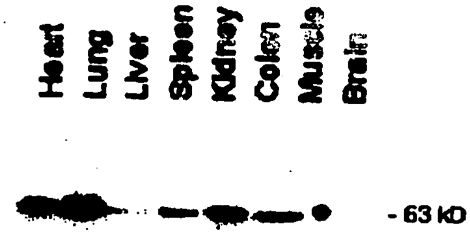


Fig. 4

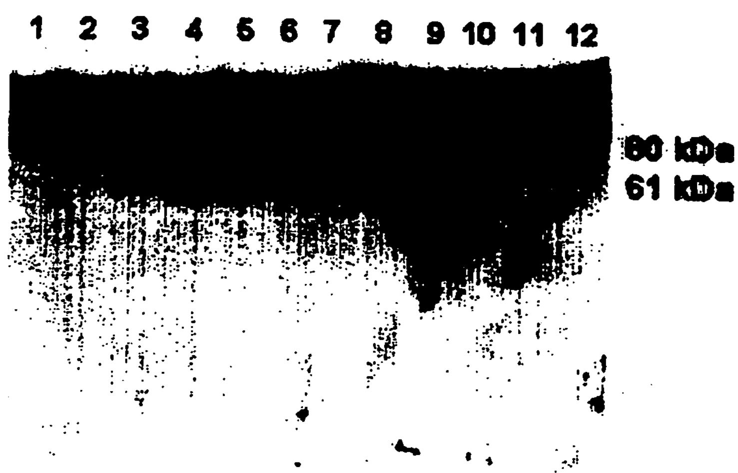


Fig. 5

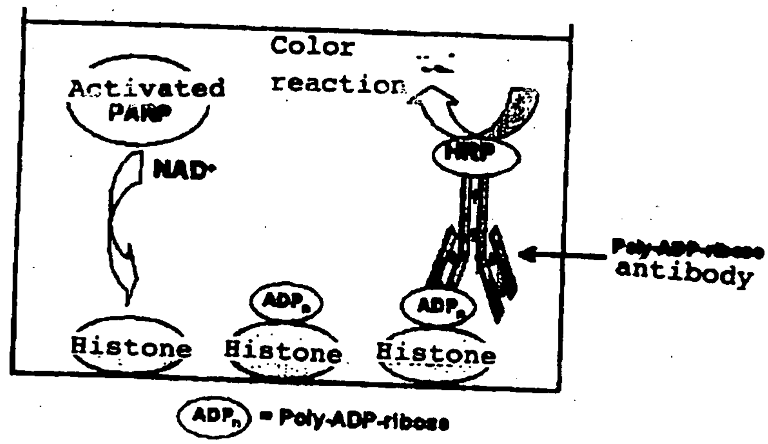

Color
reaction
Activated
PARP
NAD+
HRP
Poly-ADP-ribose
antibody
ADPn
Histone
Histone
Histone
ADPn = Poly-ADP-ribose

Fig. 6

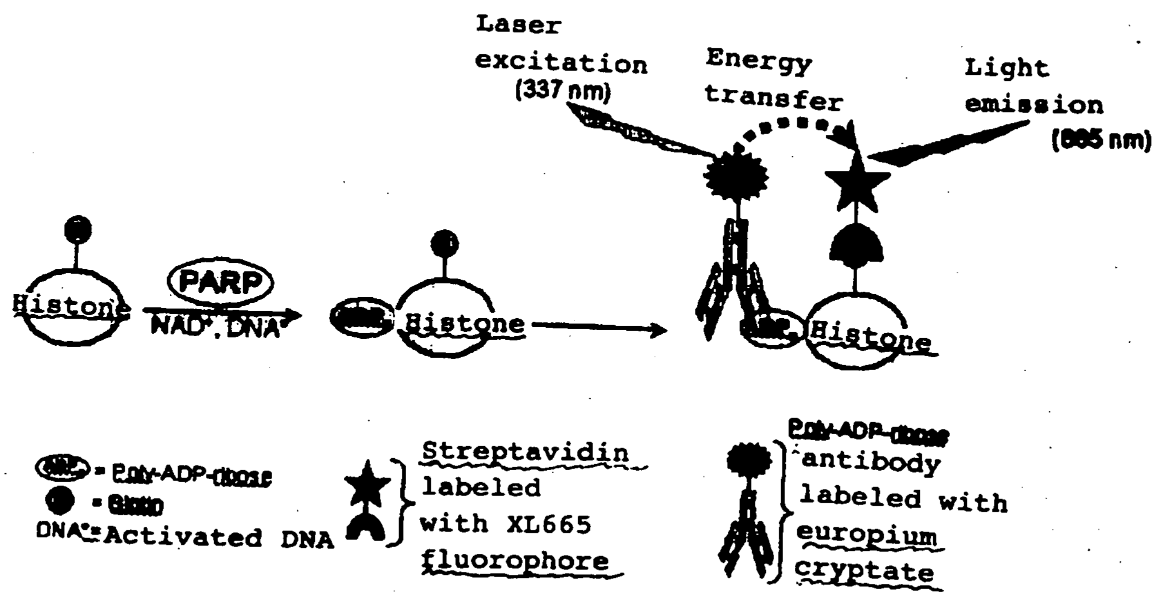

Laser
excitation
(337 nm)
Energy
transfer
Light
emission
(665 nm)
Histone
PARP
NAD+, DNA*
Histone
Histone
= Poly-ADP-ribose
= Biotin
DNA* = Activated DNA
Streptavidin
labeled
with XL665
fluorophore
Poly-ADP-ribose
antibody
labeled with
europium
cryptate

Fig. 7